Final Report on Key Comparison CCM.P-K5 in Differential Pressure from 1 Pa to 1000 Pa

A.P. Miiller[1], G. Cignolo[2], M.P. Fitzgerald[3], and M.P. Perkin[4]

January 22, 2002

ABSTRACT

This report describes a CCM key comparison of low differential-pressure standards at four National Metrology Institutes (NMIs) that was carried out during the period July 1998 to May 1999 in order to determine their degrees of equivalence at pressures in the range 1 Pa to 1000 Pa. The differential pressures were superimposed on a line pressure of nominally 100 kPa. The primary standards, which represent two principal measurement methods, included three liquid-column manometers and one double pressure balance. The transfer standard package consisted of four high-precision pressure transducers; two capacitance diaphragm gauges to provide high resolution at low differential pressures, and two resonant silicon gauges to provide the required calibration stability. Two nominally identical transfer packages were used to reduce the time required for the measurements, with Package A being circulated among laboratories in the European region (IMGC and NPL-UK) and Package B being circulated through the Asia-Pacific region (MSL-NZ). The results obtained with different transfer packages were normalized by using data obtained from simultaneous calibrations of the two packages at the pilot laboratory (NIST). The degrees of equivalence of the measurement standards were determined in two ways, deviations from key comparison reference values and pairwise differences between these deviations. The differential-pressure standards of the four participating NMIs were generally found to be equivalent and the results revealed no significant relative bias between the two principal methods tested by the comparison.

[1] NIST: National Institute of Standards and Technology, Gaithersburg, MD 20899 USA
[2] IMGC-CNR: Istituto di Metrologia "G. Colonnetti" of CNR, Torino, 10135 Italy
[3] MSL-NZ: Measurement Standards Laboratory, Lower Hutt, New Zealand
[4] NPL-UK: National Physical Laboratory, Teddington, Middlesex United Kingdom TW110LW

CONTENTS

1. INTRODUCTION

In May 1996 the Comité Consultatif pour la Masse et les grandeurs apparentéès[5] (CCM) approved proposals by the pressure working groups that identified six comparisons in pressure, the relevant ranges, the transfer standards to be used, and the pilot laboratories. The objective of these comparisons, which were seen as a scientific exercise, was to ascertain the relative performance of primary pressure standards developed at selected National Metrology Institutes (NMIs). However, with the signing of the Mutual Recognition Arrangement (MRA) [1] by NMIs of Member States of the Metre Convention in October 1999, it was agreed that the six comparisons would serve as *key comparisons* as provided for in the MRA. A major objective of the MRA is to establish the degree of equivalence of national measurement standards maintained by NMIs through key comparisons that test principal measurement methods in the field.

One of the six key comparisons was in differential pressure covering the range 1 Pa to 1000 Pa, which it was agreed would be piloted by the National Institute of Standards and Technology (NIST) using high-precision pressure transducers as transfer standards. The participants in the comparison were given the option of extending the range down to 0.1 Pa and up to 10,000 Pa, although extensions to the range would not necessarily be included in the BIPM (Bureau International des Poids et Mesures) database.

During years leading up to the start of the comparison, different types of pressure transducers were tested and evaluated at the pilot laboratory to find a transducer with sufficient pressure resolution, long-term calibration stability, and resistance to over-pressure and mechanical shock that could be used as the transfer standard. It was found that no one type of transducer could satisfy all requirements but rather a combination of two types, capacitance diaphragm gauges to provide high resolution and resonant silicon gauges to provide calibration stability.

This report summarizes the calibrations of the transfer standards carried out at four[6] NMIs during the period July 1998 and May 1999. Two nominally identical transfer standard packages were used in this comparison to reduce the time required to complete all measurements, with Package A being circulated among laboratories in the European region (IMGC and NPL-UK) and Package B being circulated through the Asia-Pacific region (MSL-NZ). Results from the two regions were normalized by using data obtained during simultaneous calibrations of the two packages at the pilot laboratory.

The following sections provide brief descriptions of the primary standards, the design and construction of the transfer standard packages, the organization and chronology of the comparison, and the general calibration procedure required by the protocol. Methods for reduction and analysis of the calibration data were chosen to provide, as much as possible, uniform treatment of the results from individual laboratories, whether they were the pilot or another participant laboratory.

2. PRIMARY STANDARDS

The principal measurement methods tested by this comparison involved two types of primary standards: a double pressure balance, which is a pressure generator; and liquid-column manometers, which are pressure measurers. One participant (MSL-NZ) used a double pressure balance as the primary standard. The remaining participants used different types of manometers in which liquid-column heights were measured either by laser interferometry (IMGC and NPL-UK) or by ultrasonic interferometry (NIST).

2.1. DIFFERENTIAL PRESSURE PRIMARY STANDARD AT THE MSL-NZ

The differential pressure primary standard at the MSL-NZ is based on two nominally identical piston-cylinder assemblies (CEC type 6-201, effective area ~ 80.6 mm^2) mounted in a common base [2]. The load on each pressure balance is first adjusted so that the pressure difference between the two instruments is less than 0.1 Pa, when generating an absolute pressure near 101 kPa. Gauge differential pressures in the

---

[5] Consultative Committee for Mass and related quantities.

[6] A fifth NMI, the Commonwealth Scientific and Industrial Research Organization (CSIRO), also took part in CCM.P-K5 initially but submitted only provisional results due to a lack of experience with their new differential-pressure standard (Casella water manometer). Their results were withdrawn after circulation of the Draft A report with the agreement of all participants.

range 1 Pa to 1000 Pa are then generated by adding small masses (16 mg to 8000 mg) to one or both pressure balances.

2.2. MM1 Laser Interferometer Manometer at the IMGC

The MM1 low-range mercury manometer is fabricated from a stainless steel block inside which two arms of a U-tube are machined with a bore and vertical length of 60 mm and 120 mm, respectively. Each arm contains a column of mercury which, at zero differential pressure, is 30 mm in height. Relative vertical displacements of the mercury menisci are measured by means of a laser interferometer, whose beams are focused on and directly reflected from the free mercury surface at the center of each mercury column using lenses mounted on floats in a cat's-eye arrangement. The manometer is not provided with active temperature control or with devices for automatic stabilization of the pressures to be measured. The MM1 can measure differential or absolute pressures up to 5 kPa. A full description and discussion of uncertainties is given in reference [3].

2.3. Laser Interferometer Manometer at the NPL-UK

The NPL low differential pressure standard is a silicon-oil manometer, capable of operating at pressures up to 1.5 kPa. It consists of two glass tubes, 100 mm in diameter, clamped between a stainless steel base plate and circular lids. A slot in the base plate connects the tubes together. The positions of the liquid surfaces are measured interferometrically with cat's-eye retro-reflecting optics which enable variations in curvature of the oil surfaces - which tend to occur as they are moved at speeds up to 0.1 mm/s - to be robustly tolerated. The liquid displacement is relatively short (±70 mm) and the cat's-eyes incorporate twin lenses mounted above the manometer's columns providing servo-controlled focusing. The oil surfaces do not reflect well at optical wavelengths - about 4% - and the interferometers are therefore designed to accommodate very large attenuation of the measurement signals. The manometer assembly is mounted in a Perspex tank, which allows temperature-controlled water to flow over the pressurized components to minimize temperature gradients in the oil and temperature-induced instabilities in the gas pressure etcetera.

2.4. Ultrasonic Interferometer Manometer at the NIST

The low differential-pressure primary standard at the NIST is a mercury Ultrasonic Interferometer Manometer (UIM), capable of operating at differential pressures up to 10 kPa [4 - 6] with reference or line pressures up to 200 kPa. The unique feature of the manometer is that changes in height of the mercury columns are determined by an ultrasonic technique. A transducer at the bottom of each liquid column generates a pulse of ultrasound (~10 MHz) that propagates vertically up the column, is reflected from the liquid-gas interface, and returns to be detected by the transducer. The length of the column, which is proportional to the change in phase of the returned signal, is determined from the phase change and the velocity of the ultrasound in mercury [7]. The manometer employs a "W" or three-column design to correct for possible tilt, large-diameter (75 mm) liquid surfaces to minimize capillary effects, thermal shields to stabilize the temperature and minimize its gradients, and high-vacuum techniques to minimize leaks and pressure gradients. An active pressure control system is used to attenuate temperature-induced pressure fluctuations to less than 10 mPa at nominal reference pressures of 100 kPa.

3. TRANSFER STANDARDS

On the basis of earlier comprehensive reviews of pressure transducer performance [8, 9], two types of gauges were selected as the transfer standards, namely, resonant silicon gauges (RSGs) for their good long-term stability and capacitance diaphragm gauges (CDGs) for their high-resolution. The RSGs are a new type of MEMS (MicroElectroMechanical Systems) sensor that have excellent calibration stability, are resistant to mechanical shock, and are only moderately susceptible to overpressure although they are rather sensitive to tilt (~ 0.4 Pa/mrad). However they lack sufficient pressure resolution to cover the entire range of the comparison. The CDGs have superior pressure resolution and, because of their all-metal

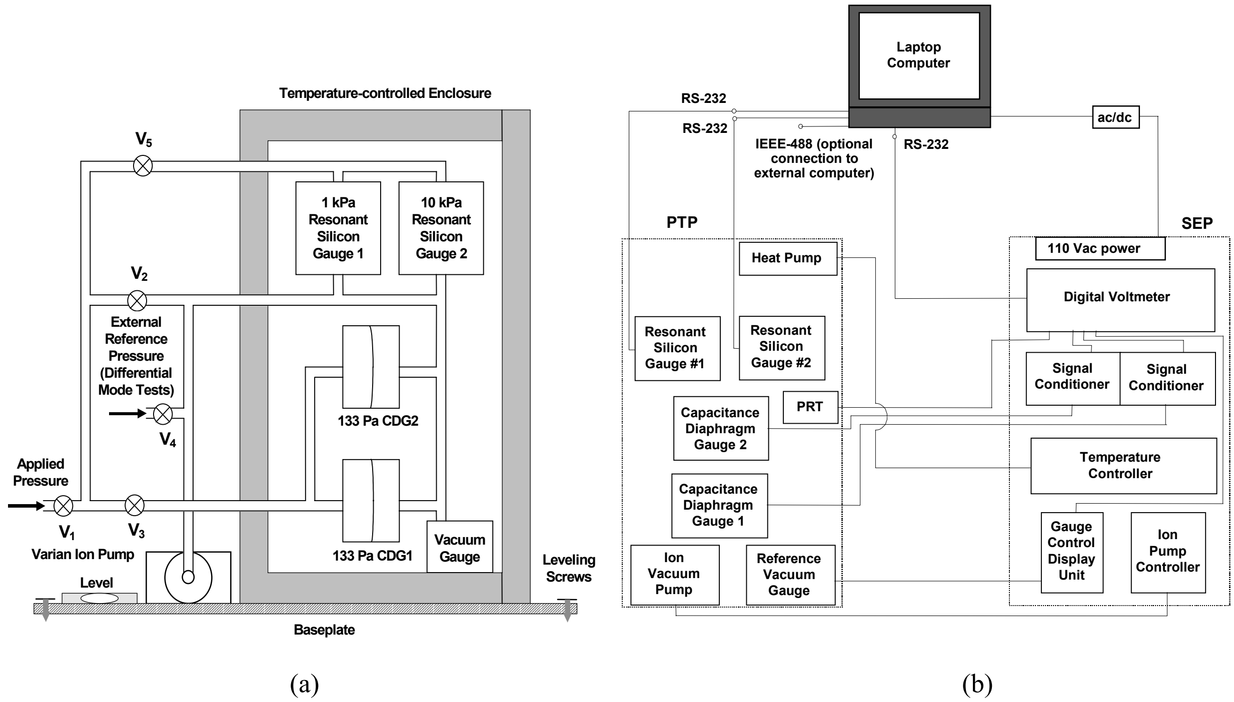


Figure 1. Schematic diagram of (a) the pressure transducer package (PTP) and (b) the electrical connections between the PTP, the support electronics package (SEP) and the laptop computer.

construction, are rugged and resistant to over-pressure but lack the desired calibration stability required by the comparison. The solution was to develop a transfer standard package using both types of gauges, two CDGs to provide redundancy and high resolution at low pressures, and two RSGs to provide redundancy and excellent calibration stability. Good calibration stability was accomplished over the entire pressure range by re-scaling the CDG response to that of the RSGs at an overlapping pressure.

The two RSGs selected for the comparison had full-scale ranges of 1000 Pa and 10,000 Pa and were combined with two CDGs each with a full-scale range of 133 Pa. Since the RSGs were available only as differential units, the decision was made to use differential CDGs as well and to use an ion vacuum pump to provide the near-vacuum reference pressure required for a companion key comparison in absolute pressure (CCM.P-K4) covering the same pressure range. This configuration gave the transfer standard the flexibility of being used for either absolute or differential mode measurements.

The transfer standard package consisted of three parts, a pressure transducer package (PTP), a support electronics package (SEP), and a laptop computer (see Figure 1). The PTP included four differential transducers housed in a temperature-controlled enclosure, a calibrated 100-ohm platinum resistance thermometer (PRT) to measure the interior temperature of the enclosure, and an ion vacuum pump and reference-pressure vacuum gauge for the absolute mode calibrations. All-metal plumbing was used throughout the PTP including five metal bellows-sealed valves and metal bellows connections to each transducer to minimize mechanical strain. The valves included external isolation valves V_1 and V_4, internal isolation valves for the CDGs (V_3) and RSGs (V_5), and an internal bypass valve V_2 between the pressure and reference side of the gauges. The gauges and internal plumbing were maintained under vacuum during shipment or storage, but with all internal valves open to avoid overpressurization of the gauges in the event of a leak to atmosphere. The tilt orientation of the PTP during calibration of the RSGs was monitored by means of sensitive bubble levels mounted on the PTP base plate and any observed changes were corrected using the leveling screws.

The SEP included a temperature controller for the transducer enclosure, signal-conditioning electronics for the CDGs, a controller for the ion vacuum pump, and a digital voltmeter (DVM) for digitizing analog signals from the CDGs, the calibrated PRT, and the reference vacuum gauge. The enclosure temperature was controlled by means of a heat pump and a Wheatstone bridge mounted inside the enclosure where the bridge included an uncalibrated PRT and an adjustable resistor in two of its arms (not shown in Figure 1b).

A laptop computer was used for controlling the acquisition of data from the RSGs and the DVM during calibration. The time required to obtain one set of readings was approximately 55 seconds. Because of their accuracy (~ 1 part in 10^4), the readings of the RSGs were multiplied by a scale factor before display and storage on the laptop computer in order to increase the level of confidentiality for the pilot laboratory data. The RSG data submitted by the participants were multiplied by 1/(scale factor) during subsequent data reduction in order to restore the original readings.

For interlaboratory shipment, the PTP and SEP (with the laptop) were mounted in commercial containers that were specially designed for vibration and shock isolation.

4. ORGANIZATION OF THE KEY COMPARISON

The present key comparison in differential pressure (CCM.P-K5) was organized in conjunction with another key comparison in absolute pressure (CCM.P-K4) covering the same range in order to minimize the time required for completion of both comparisons. Two nominally identical transfer standard packages were developed for use in either absolute or differential mode. Transfer standard package A was circulated through the European region while transfer standard package B was circulated through the Asian-Pacific region with calibrations at the pilot laboratory (NIST) at the start and end of the comparison.

4.1. CHRONOLOGY OF THE MEASUREMENTS

Table 1 presents the actual chronology of calibrations during the measurement phase of both comparisons. The start and end dates refer to the measurement time period during which calibration data was taken at each NMI. The total time required to complete the measurement phase of both absolute and differential pressure comparisons was eighteen months, which is approximately three months longer than originally projected due to unanticipated problems (see next section).

The sequence of calibrations for the **key comparison in differential pressure** included simultaneous calibrations of the two transfer standard packages during the second and third calibration cycles at the pilot laboratory (NIST #2 and NIST #3) as follows:

Package A: **NIST #2** => NPL-UK => IMGC => **NIST #3**

Package B: **NIST #2** => MSL-NZ => CSIRO => **NIST #3**

4.2. PROBLEMS DURING THE COMPARISONS

There were several problems during the course of the comparisons, ranging from instrument malfunction during calibration to severe damage to the transfer standard package during shipment between NMIs. In most cases instrument failures were remedied by replacement with an equivalent unit. However a potentially more serious instrument failure was the rare but intermittent malfunction of one of the 10 kPa RSGs, first observed during the initial evaluation of these gauges. The manufacturer was unable to diagnose the problem but did send a new processor board as a backup. At the start of the first calibration of Package B (NIST #2) the offending RSG began to malfunction again and the processor board was replaced. Unexpectedly during calibrations at the IMGC, the 10 kPa RSG in Package A also began to exhibit the same behavior after completing only one run of the absolute mode calibrations but began operating normally again when used for the differential mode calibrations. The gauge continued to operate normally after Package A was shipped to the pilot laboratory for its third and final calibration (NIST #3).

Very rough handling of the transfer standard containers during shipment caused the most severe problems and contributed to significant delays in completing the comparisons. When Package B arrived at the NPL-I, the gauges in the PTP were found to be at atmosphere due to the rupture of a metal bellows inside the thermal enclosure. A replacement bellows was fabricated at the NIST and sent to the NPL-I to

Table 1. Chronology of measurements during two key comparisons, one in differential pressure (in bold) the other in absolute pressure.

NMI	Transfer Std Package	Calibration Start Date	Calibration End Date	Calibration Mode
		European Region		
NIST #1	A	March 17, 1998	April 3, 1998	Absolute
PTB†	A	May 28, 1998	June 4, 1998	Absolute
NIST #2	**A**	July 8, 1998 **July 30, 1998**	July 24, 1998 **Aug 5, 1998**	Absolute **Differential**
NPL-UK	**A**	**Oct 29, 1998** Nov 17, 1998	**Nov 4, 1998** Nov 25, 1998	**Differential** Absolute
IMGC	**A**	Jan 19, 1999 **Feb 18, 1999**	Feb 8, 1999 **Feb 23, 1999**	Absolute **Differential**
NIST #3	**A**	April 23, 1999 **May 11, 1999**	May 6, 1999 **May 17,1999**	Absolute **Differential**
		Asia-Pacific Region		
NIST #2	**B**	July 8, 1998 **July 30, 1998**	July 24, 1998 **Aug 5, 1998**	Absolute **Differential**
MSL-NZ	**B**	**Oct 30, 1998**	**Nov 5, 1998**	**Differential**
NPL-I††	B	Jan 1, 1999	Jan 14, 1999	Absolute
CSIRO‡	**B**	Feb 24, 1999 **March 11, 1999**	March 3, 1999 **March 18, 1999**	Absolute **Differential**
NIST #3	**B**	April 23, 1999 **May 11, 1999**	May 6, 1999 **May 17,1999**	Absolute **Differential**
KRISS‡‡	B	June 15, 1999	June 22, 1999	Absolute
NIST #4	B	Aug 23, 1999	Sept 10, 1999	Absolute

† Physikalisch-Technische Bundesanstalt,
†† National Physical Laboratory – India,
‡ Commonwealth Scientific and Industrial Research Organization; results withdrawn from comparison CCM.P-K5
‡‡ Korean Research Institute of Science and Standards

enable the repair. The most serious damage to Package B however occurred during shipment from the CSIRO to the pilot laboratory and was consistent with penetration of the PTP container by a forklift truck. The force of the impact was sufficient to dismount both CDGs and rupture several metal bellows inside the thermal enclosure. Remarkably after Package B was repaired and re-calibrated (NIST #3), the gauges did not show any unusual shifts in their calibration (see Section 6.2).

5. GENERAL CALIBRATION PROCEDURE

The general procedure for the key comparison required that each laboratory calibrate the transfer standard (with nitrogen gas) at the following nominal differential pressures in ascending order: 1 Pa, 3 Pa, 10 Pa, 30 Pa, 100 Pa, 300 Pa, and 1000 Pa. The reference or baseline pressure was required to be in the range 95 kPa to 105 kPa. The actual differential pressures realized at the transfer standard gauges by the participant's pressure standards were to be within 2 parts in 100 of the target pressures. Optional calibration data[7] could also be taken at 0.1 Pa, 0.3 Pa, and at 3000 Pa, 10,000 Pa.

A total of **five calibration runs** were required, with each run taken on a different day. Within a calibration run, **five repeat sets** of pressure and temperature readings of the transfer standard and primary standard were required at each target pressure. At the beginning of each calibration run, **ten repeat sets** of zero-pressure readings for the transfer standard gauges were required to be taken with the PTP isolated from the participant's calibration system (valves V_1 and V_4 closed) and with internal isolation valves V_3 and V_5 and bypass valve V_2 open. These data were needed to correct calibration data obtained with liquid-

[7] Optional data, which were taken by IMGC and by NIST at 3000 Pa, are not included in this report.

column manometers for zero-pressure offsets. An additional ten repeat sets of zero-pressure readings were to be taken at the end of each run in order to monitor zero drift in the four transducers during calibration. The calibration procedure also included the option of recording **five sets** of zero-offset readings for the gauges just prior to establishing each target pressure. These readings, which were taken with the external and internal isolation valves of PTP open and bypass valve V_2 closed, were needed to correct zero offsets in calibration data obtained with the double pressure balance.

The format for reporting calibration data followed the measurement sequence dictated by the data acquisition software. The sequence for each set of associated readings of the transfer standard and the participant's primary standard was:

Set No. p_{CDG1} p_{CDG2} p_{RSG1} p_{RSG2} p_{REF} t_{PRT} P_{STD} t_{STD}

where the meaning of subscripts for pressures p (gauges), P (primary standard) and temperatures t are self-explanatory. All calibration data were transmitted to the pilot laboratory in the form of spreadsheet files, which greatly facilitated the analyses of a rather voluminous amount of data.

6. REDUCTION AND ANALYSIS OF THE REPORTED DATA

The reduction and analysis of the key comparison data required that several factors be addressed. These included zero-pressure offsets (Section 6.1), deviations of the actual pressures realized from the target pressures (Section 6.2), relatively large calibration shifts in the capacitance diaphragm gauges (Section 6.3), and normalization of data from two different transfer standard packages (Section 6.4). Methods for estimating uncertainties (Sections 6.5 and 6.6) and for evaluating degrees of equivalence (Section 6.7) are also described.

6.1. CORRECTIONS FOR ZERO-PRESSURE OFFSETS

The first step in reducing the comparison data is to correct the readings of each gauge i for its zero-pressure offset. The index i is equal to either 1 or 2 and refers to either, CDG1 and CDG2, or RSG1 and RSG2 (see Figure 1). At a given target pressure during calibration run k, the corrected reading of gauge i for repeat set l is given by:

$$p_{ikl} = p_{Gikl} - \langle p_{Gik0} \rangle_{10} \qquad \text{for liquid-column manometer data} \qquad (1)$$

and

$$p_{ikl} = p_{Gikl} - \left[\langle p_{Gik0} \rangle_{5a} + \langle p_{Gik0} \rangle_{5b} \right] / 2 \qquad \text{for double pressure balance data} \qquad (2)$$

where p_{Gikl} is the uncorrected gauge reading, $\langle p_{Gik0} \rangle_{10}$ is the mean of 10 zero-pressure readings taken just prior to the start of calibration run k and, $\langle p_{Gik0} \rangle_{5a}$ and $\langle p_{Gik0} \rangle_{5b}$ are the means of 5 zero-offset readings, one taken just before and the second just after realizing each target pressure, respectively.

6.2. CALCULATION OF CALIBRATION RATIOS

The transfer standard gauges are nominally linear devices and so the ratio of transfer standard reading to primary standard reading will be essentially independent of pressure for a range of pressures about each target value. Once calculated these calibration ratios are used to correct the gauge readings for deviations of the primary standard from the target pressure and so they form the basis for the comparison of measurement standards from different NMIs.

At each target pressure during calibration run k the mean ratio of 5 sets of repeat readings of transfer standard gauge i and primary standard j is given by

$$a_{ijk} = \frac{1}{5} \sum_{l=1}^{5} \frac{p_{ikl}}{P_{jkl}} \qquad (3)$$

where p_{ikl} and P_{jkl} are the "simultaneous" readings of the gauge and primary standard, respectively. The mean of the a_{ijk} for 5 calibration runs defines a *calibration ratio* given by

$$a_{ij} = \frac{1}{5}\sum_{k=1}^{5} a_{ijk} = \frac{1}{25}\sum_{k=1}^{5}\sum_{l=1}^{5}\frac{p_{ikl}}{P_{jkl}} \tag{4}$$

The calibration ratio, if expressed as

$$a_{ij} = \frac{p_i}{P_j}, \tag{5}$$

may be used to calculate a gauge reading p_i from the pressure being measured/generated[8] by primary standard j, P_j, or vice-versa.

Figures 2 and 3 present the calibration ratios for CDGs and RSGs in the two transfer standard packages as determined by two simultaneous differential-mode calibrations of the packages at NIST. The superior stability of the RSGs is clearly evident, even at 100 Pa where the long-term shifts in their response between calibrations is about a factor of 50 smaller than similar shifts exhibited by the CDGs. It is remarkable that, despite the damage sustained by Package B during shipment to the pilot laboratory for calibration NIST #3, the shifts in calibration of its gauges were not unusual nor significantly different in magnitude from calibration shifts observed for gauges in Package A.

6.3. RE-SCALING OF THE CDG READINGS

The relatively large calibration shifts of the CDGs can be reduced significantly by re-scaling their readings so they equal those of the RSGs (in the same package) at an overlapping pressure, namely 100 Pa. The readings of the two CDGs at 100 Pa could be re-scaled to a single RSG (either RSG1 or RSG2) or to the mean of two RSGs but then the re-scaled readings of CDG1 and CDG2 would not be independent as required for Youden graphical analyses in Section 7. Although arbitrary, it was decided to pair CDG1 with RSG1, and CDG2 with RSG2, when re-scaling the CDG readings.

At target pressures p_t < 100 Pa, the re-scaled readings of capacitance diaphragm gauge i may be expressed as

$$p_{CDGi}(p_t) = p_{Gi}(p_t)\left[\frac{p_{RSGi}(100)}{p_{Gi}(100)}\right] \tag{6}$$

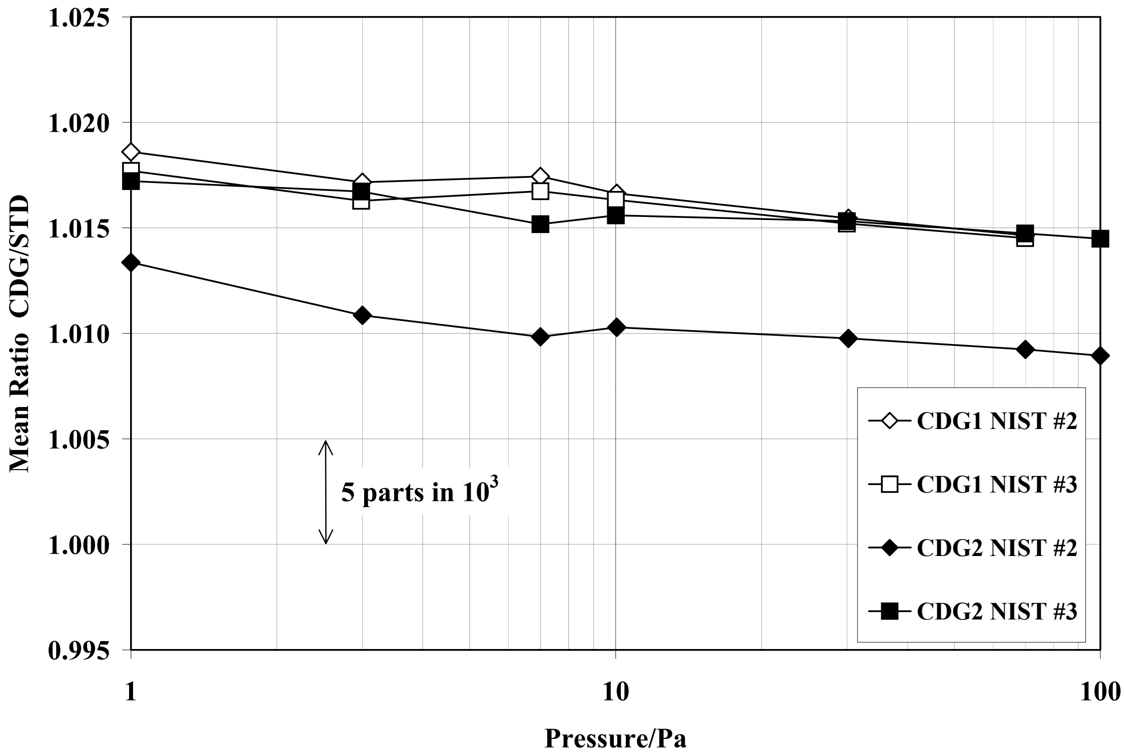
where $p_{Gi}(p_t)$ is the CDG reading before re-scaling. This equation may be re-expressed in terms of calibration ratios by means of Equation (5) as

$$a_{CDGij}(p_t) = a_{Gij}(p_t)\left[\frac{a_{RSGij}(100)}{a_{Gij}(100)}\right] \tag{7}$$

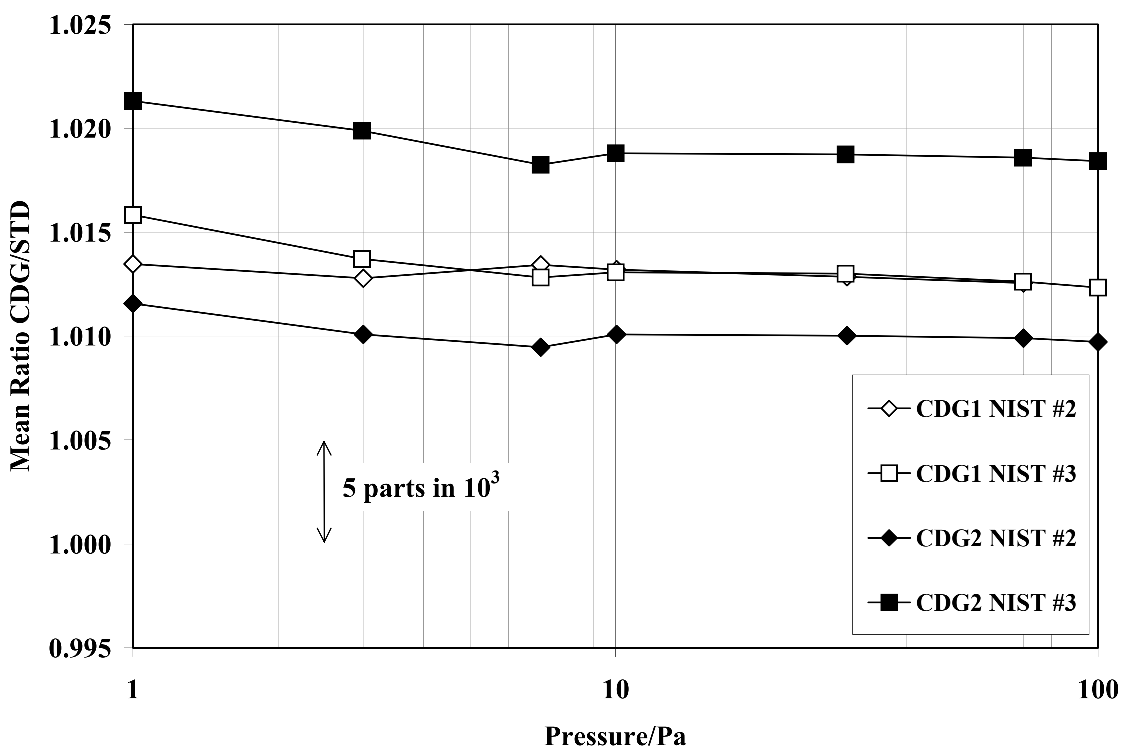
where a_{Gij} and a_{CDGij} are the respective calibration ratios for capacitance gauge i before and after re-scaling, and a_{RSGij} is the calibration ratio for resonant silicon gauge i. As Figure 5 shows, the observed shifts in the CDG ratios between successive calibrations at NIST are substantially reduced by re-scaling even though sizeable changes in linearity of the CDG response remain at lowest pressures.

In summary, the present key comparison is based on pairs of calibration ratios, a_{CDG1j} and a_{CDG2j} for pressures lower than 100 Pa, and on a_{RSG1j} and a_{RSG2j} for pressures 100 Pa up to and including 1000 Pa.

[8] The measured or generated pressure is the calculated value obtained from measurements of mercury (or oil) temperature and column-height changes in manometers, or from measurements of the added masses and the effective area and temperature of the piston/cylinder assemblies in double pressure balances, respectively.

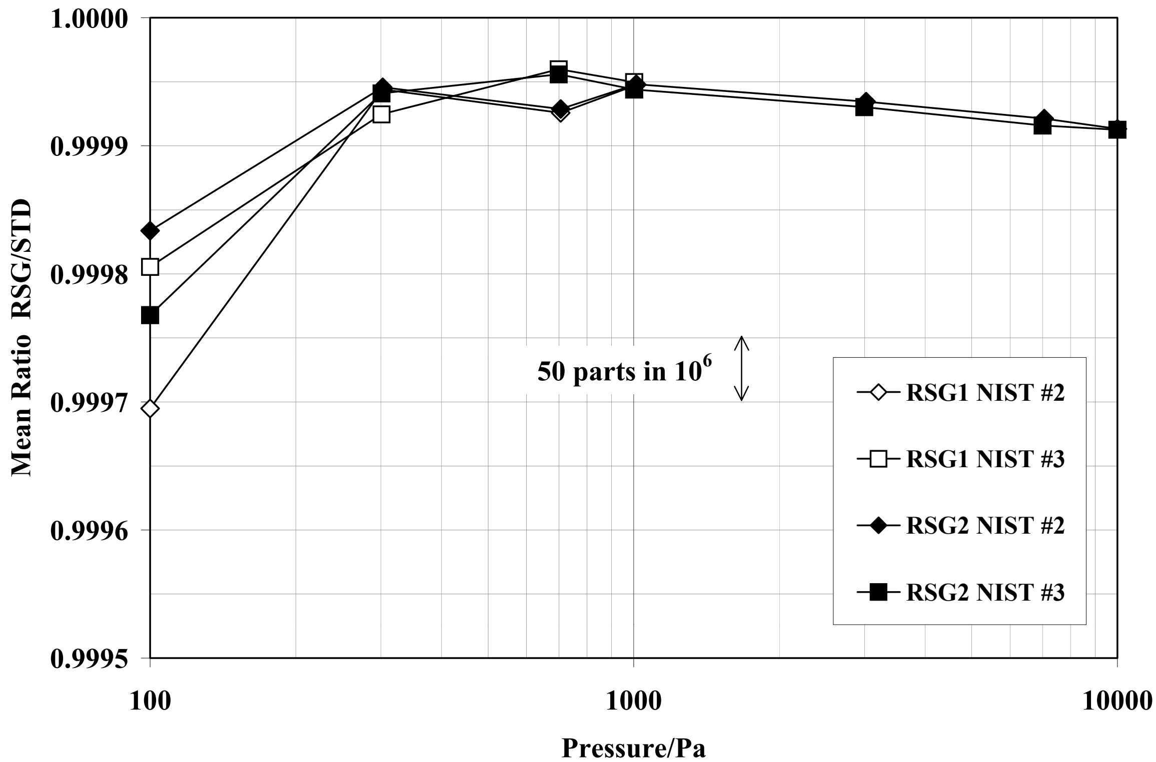


Transfer Package A

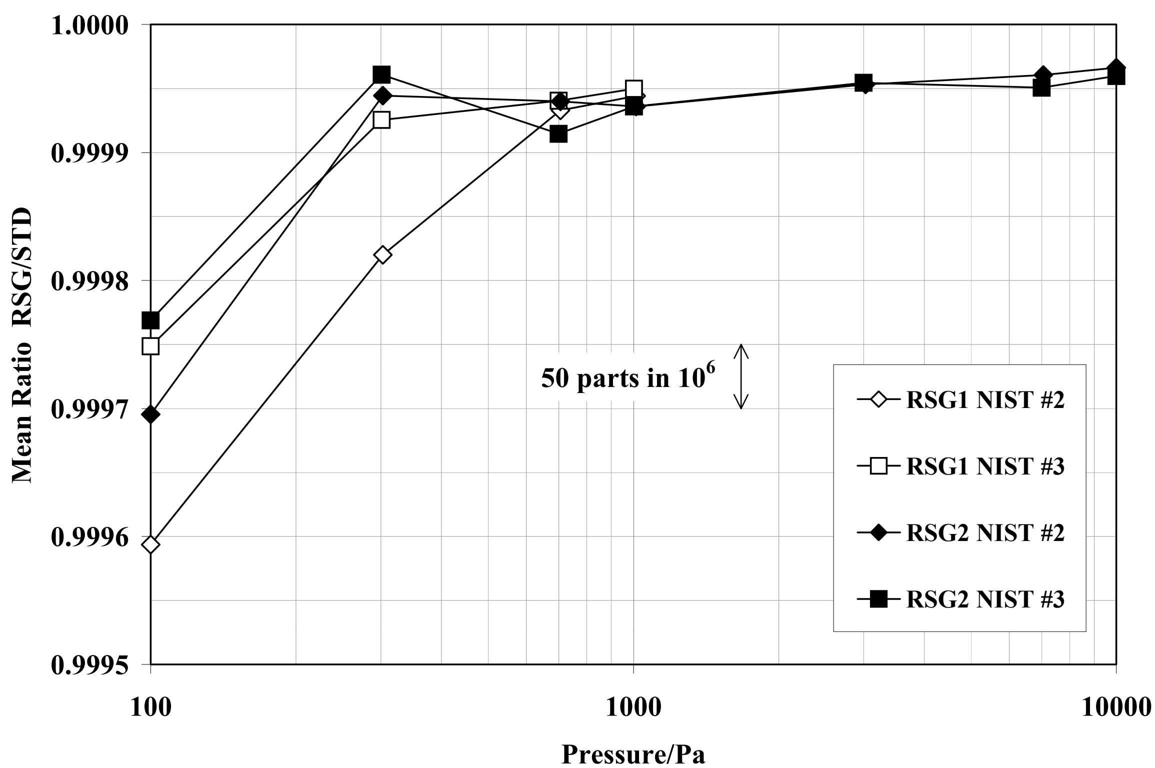


Transfer Package B

Figure 2. Calibration ratios for CDGs in Transfer Package A (upper) and Transfer Package B (lower) as a function of pressure.

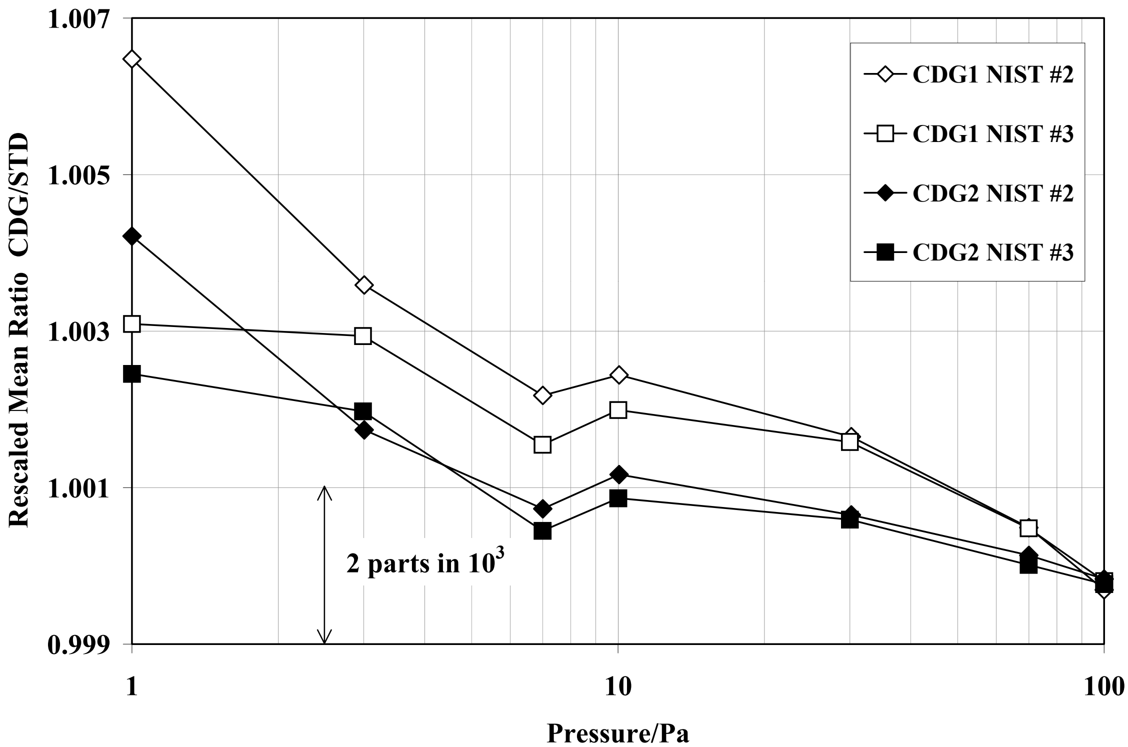


Transfer Package A

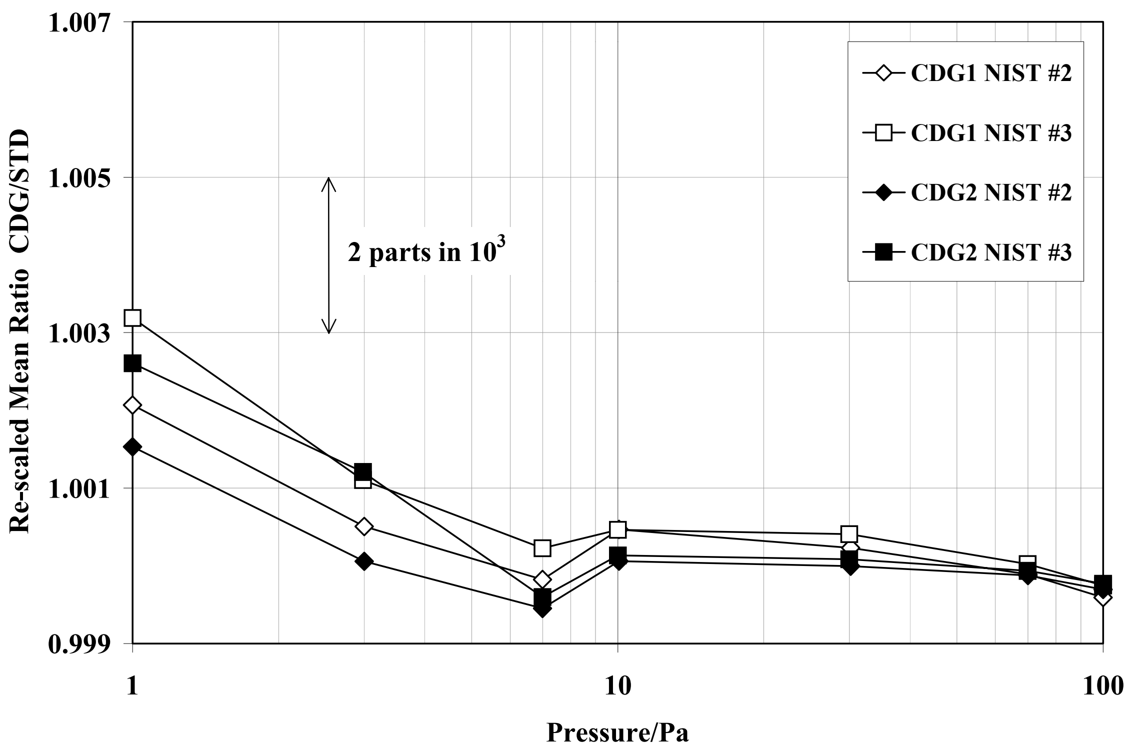


Transfer Package B

Figure 3. Calibration ratios for RSGs in Transfer Package A (upper) and Transfer Package B (lower) as a function of pressure.



Transfer Package A



Transfer Package B

Figure 4. Calibration ratios for the CDGs after re-scaling to the RSGs, as a function of pressure.

6.4. CALCULATION OF THE PREDICTED GAUGE READINGS

Degrees of equivalence [1] of the primary standards for differential pressure can be expressed quantitatively by comparing pressure readings of the transfer standard gauges. The basic method adopted here is to use the calibration ratios to predict gauge readings that would be observed when different primary standards measure/generate pressures exactly equal to the target value. The difference in the predicted gauge readings is taken as a surrogate for the difference between "true" pressures actually realized by the different primary standards[9]. Results obtained with the two transfer standard packages are normalized by using data taken during simultaneous calibrations at the pilot laboratory.

At target pressures up to and including 1000 Pa there are two gauges ($i = 1, 2$). Thus for either package there will be two gauge readings for each pressure measured/generated by primary standard j and, according to Equation (5), these may be expressed as:

$$p_{ijA} = a_{ij} P_j \qquad \text{or} \qquad p_{ijB} = b_{ij} P_j \tag{8}$$

where a_{ij} and b_{ij} are the calibration ratios for gauges i in Packages A and B, p_{ijA} and p_{ijB} are their respective pressure readings, and P_j is the measured/generated pressure. Clearly, gauge readings from a given package could be used to compare the primary standards used for their calibration. However to compare primary standards used to calibrate different transfer packages requires that the relationship between gauge readings in the two packages be known.

The relationship between gauge readings i in Packages A and B can be determined by simultaneous calibration against primary standard j and may be expressed as

$$\frac{p_{ijA}}{a_{ij}} = \frac{p_{ijB}}{b_{ij}} = P_j \tag{9}$$

i.e., the ratio of readings of each pair of gauges i in the two packages is equal to the inverse ratio of their calibration ratios determined during the simultaneous calibration. The ratios of calibration ratios, once determined by the simultaneous calibration, could be used to convert all comparison data from Package B to equivalent data from Package A, or vice versa. However, this would in effect reference the key comparison to one package, either Package A or Package B.

Alternatively, Equation (9) can be expressed as

$$\frac{p_{ij}}{n_{iA} a_{ij}} = \frac{p_{ij}}{n_{iB} b_{ij}} = P_j \tag{10}$$

where n_{iA} and n_{iB} are coefficients that re-scale or normalize the readings of gauges i in the two packages to a common normalized gauge reading, p_{ij}, according to

$$p_{ij} = n_{iA} p_{ijA} = n_{iB} p_{ijB} \tag{11}$$

A second property needed to define the normalization coefficients is

$$p_{ij} = (p_{ijA} + p_{ijB})/2 \tag{12}$$

i.e., the normalized gauge reading is also equal to the mean reading of gauges i in Packages A and B obtained during their simultaneous calibration. The significance of Equation (10) when re-written as

$$p_{ij} = n_{iA} a_{ij} P_j = n_{iB} b_{ij} P_j \tag{13}$$

is that it predicts the same normalized gauge reading p_{ij} for a given measured/generated pressure P_j whether calibration ratios for Package A or those for Package B are used. Therefore, once the normalization coefficients have been determined, Equation (13) provides a common basis for comparing results obtained by participants using Package A with results obtained from Package B.

---

[9] The difference between "true" pressures being realized by two primary standards, when set to measure/generate a target pressure p_t, is (to a very good approximation) equal to but has the opposite sign of the difference between the measured/generated pressures when both standards realize the same "true" pressure equal to p_t.

Table 2. Normalization coefficients n_{iA} and n_{iB} for gauges in transfer standard packages A and B, respectively, where values above the dotted line refer to CDGs and those below refer to RSGs.

Package	Target Pressure Pa	From NIST #2		From NIST #3		Mean Coefficient	
		CDG1 / RSG1	CDG2 / RSG2	CDG1 / RSG1	CDG2 / RSG2	CDG1 / RSG1	CDG2 / RSG2
A	1	0.997809	0.998665	1.000048	1.000077	0.998928	0.999371
	3	0.998464	0.999161	0.999085	0.999617	0.998775	0.999389
	10	0.999016	0.999446	0.999236	0.999636	0.999126	0.999541
	30	0.999292	0.999673	0.999413	0.999746	0.999353	0.999710
	100	0.999949	0.999931	0.999972	1.000001	0.999960	0.999966
	300	0.999938	0.999999	1.000000	1.000010	0.999969	1.000005
	1000	0.999998	0.999994	1.000000	0.999996	0.999999	0.999995
B	1	1.002200	1.001339	0.999952	0.999923	1.001076	1.000631
	3	1.001541	1.000840	1.000916	1.000384	1.001228	1.000612
	10	1.000986	1.000555	1.000765	1.000364	1.000876	1.000460
	30	1.000709	1.000327	1.000588	1.000254	1.000648	1.000290
	100	1.000051	1.000069	1.000028	0.999999	1.000040	1.000034
	300	1.000062	1.000001	1.000000	0.999990	1.000031	0.999995
	1000	1.000002	1.000006	1.000000	1.000004	1.000001	1.000005

The normalization coefficients, n_{iA} and n_{iB}, were expressed in terms of calibration ratios determined from simultaneous calibration of the two transfer packages at the pilot laboratory (PL) via Equations (11) and (12):

$$n_{iA} = \frac{a_{iPL} + b_{iPL}}{2a_{iPL}} \qquad \text{and} \qquad n_{iB} = \frac{a_{iPL} + b_{iPL}}{2b_{iPL}} \tag{14}$$

Table 2 presents values for the coefficients and their mean obtained from two simultaneous calibrations, NIST #2 and #3.

Thus, the procedure is to calculate the normalized gauge readings, p_{ij}, that would be obtained when the pressure measured/generated by each primary standard j equals the target pressure, p_t, i.e., when $P_j = p_t$. The respective normalized gauge readings for Package A and for Package B are then obtained from Equation (13) as:

$$p_{ij} = n_{iA} a_{ij} p_t \qquad \text{and} \qquad p_{ij} = n_{iB} b_{ij} p_t \tag{15}$$

The results for p_{ij} from individual laboratories are presented in Section 7.1.

At each target pressure up to and including 1000 Pa, there are two values for the normalized readings (e.g., for CDG1 and CDG2, etc.) and so a mean normalized gauge reading p_{jU} was calculated as a simple arithmetic mean:

$$p_{jU} = \frac{p_{1j} + p_{2j}}{2} \tag{16}$$

where the subscript U denotes that gauge readings p_{ij} are uncorrected. For the pilot laboratory, a single value of p_{jU} was calculated as the arithmetic mean of eight values of p_{ij}^{mn}. The values of p_{ij}^{mn} were determined via Equation (15) using calibration ratios a_{ij}^{mn} or b_{ij}^{mn} obtained from two simultaneous calibrations (n = 1, 2) of two gauges (i = 1, 2) in two transfer standard packages (m = A, B).

The "true" pressures realized by the primary standards when set to measure/generate a given target pressure should, on average, closely approximate the SI value under the assumption that deviations from the SI value are randomly distributed. Therefore, it is reasonable to correct the mean normalized gauge readings so that their ensemble average is also equal to the target pressure. Thus, the corrected mean gauge readings can be expressed as

$$p_j = f_C p_{jU} = f_C \left(\frac{p_{1j} + p_{2j}}{2} \right) \tag{17}$$

where the correction factor f_C is obtained by setting the ensemble average of the p_j values from the calibrations at individual laboratories equal to the target pressure (see Section A1 of the Appendix). The resultant values for f_C are very nearly equal to one (see Table A1). The results for p_j from individual laboratories are presented in Section 7.2.

Implicit in the above analysis is the assumption that response functions of the transfer standard gauges do not change during the comparison. Of course this is not true (see Figures 3 and 4) since the long-term shifts in gauge response as well as other sources will contribute uncertainty to the normalized gauge readings p_{ij} and ultimately to the corrected mean gauge readings p_j.

6.5. ESTIMATES OF UNCERTAINTIES IN THE NORMALIZED GAUGE READINGS

The combined uncertainty[10] in the normalized gauge readings calculated using Equation (15) may be estimated from the root-sum-square of three component uncertainties [10, 11],

$$u_c(p_{ij}) = \sqrt{u_{std}^2(p_{ij}) + u_{rdm}^2(p_{ij}) + u_{lts}^2(p_{ij})} \tag{18}$$

where $u_{std}(p_{ij})$ is the uncertainty in p_{ij} due to systematic effects in primary standard j, $u_{rdm}(p_{ij})$ is the uncertainty in p_{ij} due to the combined effect of short-term random errors of transfer standard gauge i and primary standard j during calibration, and $u_{lts}(p_{ij})$ is the uncertainty arising from long-term shifts in the response function of gauge i during the course of the comparison.

Table 3 and Figure 5 present the estimated relative uncertainties in pressure arising from systematic effects in the primary standards, as stated by the participants for target pressures used in the comparison. Such estimates usually involve both Type A and Type B evaluations.

The relative uncertainty in p_{ij} due to short-term random effects during calibration can be estimated from the corresponding uncertainties in the calibration ratios via Equation (15):

$$\frac{u_{rdm}(p_{ij})}{p_{ij}} = \frac{u_{rdm}(a_{ij})}{a_{ij}} \tag{19}$$

Similarly the relative uncertainty in p_{ij} due to long-term shifts in gauge response between calibrations is given by

$$\frac{u_{lts}(p_{ij})}{p_{ij}} = \frac{u_{lts}(a'_{ij})}{a'_{ij}} \tag{20}$$

where $a'_{ij} = n_{iA} a_{ij}$ is the normalized calibration ratio for gauge i in Package A. It is understood that the above equations involving a_{ij} or n_{iA} apply equally well to b_{ij} and n_{iB}.

---

[10] Uncertainty refers to standard uncertainty unless noted otherwise.

Table 3. Relative standard uncertainties, as stated by the participants, due to systematic effects in their primary standards. Not all digits are significant but are retained for calculation of final results.

Target Pressure Pa	$100 \times u_{std}(p_t)/p_t$			
	Pressure Balance	Liquid-Column Manometers		
	MSL-NZ	IMGC	NIST	NPL-UK
1	0.4014	0.6672	0.1221	
3	0.1347	0.2339	0.0775	0.0550
10	0.0414	0.0822	0.0300	0.0200
30	0.0147	0.0389	0.0100	0.0100
100	0.0054	0.0237	0.0030	0.0065
300	0.0027	0.0084	0.0010	0.0055
1000	0.0018	0.0030	0.0004	0.0052

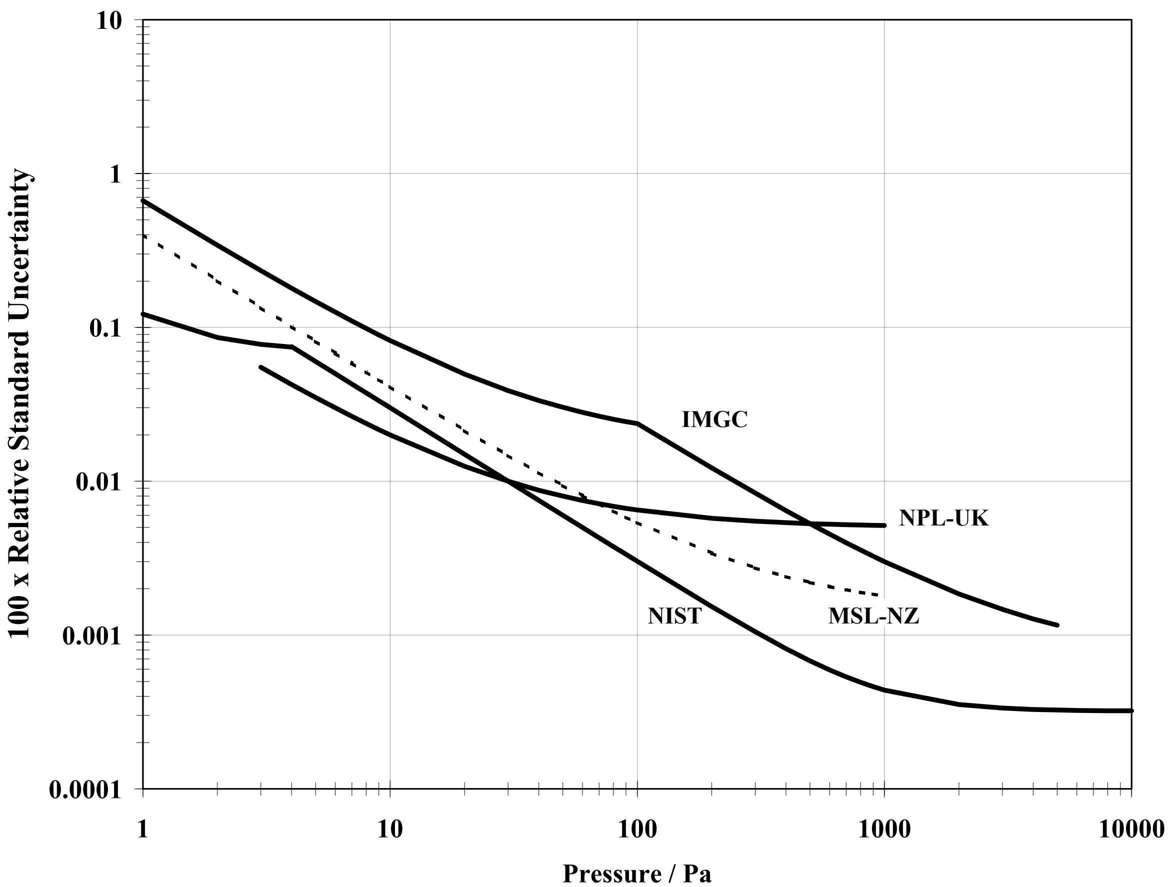


Figure 5. Relative uncertainty due to systematic effects in primary standards at the participating laboratories as a function of pressure. The heavy lines identify uncertainties associated with liquid-column manometers and the medium dashed lines identify uncertainties for the double pressure balance.

The short-term random uncertainty in a calibration ratio, a_{ij}, as given by Equation (4), may be estimated by a Type A evaluation in one of two ways, either as

$$u_{rdm}(a_{ij}) = \sigma_{ijkl} \big/ \sqrt{25} \tag{21}$$

where σ_{ijkl} is the standard deviation of 25 values of p_{ikl} / P_{jkl} about their mean a_{ij} OR

$$u_{rdm}(a_{ij}) = \sigma_{ijk} \big/ \sqrt{5} \tag{22}$$

where σ_{ijk} is the standard deviation of five values of the means, a_{ijk}, about their mean a_{ij}. Equation (21) is appropriate only if the "true" means for the five runs are equal, that is, they refer to the same parent population of observations. If not, it suggests that the run-to-run (day-to-day) variability is dominant in which case Equation (22) should be used to calculate short-term random uncertainties.

An analysis of variance (ANOVA) was performed on the five sets of five pressure ratios for each gauge at each target pressure to test the hypothesis that the "true" means from five calibration runs are equal. For nearly all comparison data, the hypothesis was rejected with a less than 5 % chance that the means are actually equal. For the few remaining data, the hypothesis was rejected with a somewhat larger chance for error (~50 % or less). Therefore, Equation (22) was used to estimate the short-term random uncertainties, $u_{rdm}(a_{ij})$ or $u_{rdm}(b_{ij})$, which are given in columns seven and eight of Table 5 in Section 7.1. The short-term random uncertainties in the re-scaled calibration ratios obtained via Equation (7) were estimated as the root-sum-square of component uncertainties arising from random effects in $a_{Gij}(p_t)$, $a_{Gij}(100)$, and $a_{RSGij}(100)$, each evaluated using Equation (22).

The ANOVA tests confirm an expected result, and that is, random effects due to operational differences between five calibration runs performed on five different days are generally not the same as, and are larger than, random effects associated with five repeat readings taken during a period of five minutes. Possible sources of the short-term day-to-day variability include differences in zero drifts of the gauges, differences in achieving stable target pressures, etc. This variability was assumed to be random and uncorrelated for each pair of gauges.

Long-term shifts in gauge response are often one of the largest component uncertainties, particularly for CDGs, yet most difficult to evaluate. The reasons for this are often twofold: (a) the number of repeat calibrations against the same standard is limited for practical reasons and (b) the effect of transportation between laboratories (rough handling, etc.) is unknown. Earlier studies at the pilot laboratory [8, 9] have shown that changes in response functions of CDGs and RSGs between calibrations generally do not occur as a monotonic drift with time (over intervals of months to years) but rather as shifts that are essentially random in both sign and magnitude. Furthermore, the earlier studies showed that, at least for low range CDGs, the magnitude of the shifts was on average about a factor of two larger for gauges transported between laboratories than for gauges maintained at the pilot laboratory.

In the present comparison, the observed shifts in gauge response between calibrations at the pilot laboratory (see Figures 2 and 3) are consistent with the earlier studies, i.e., there is little evidence that the calibration shifts are systematic (even for the RSGs) but rather they appear more random in character. Therefore, the observed variability in gauge response was assumed to be purely random but, because the statistical sample of pilot laboratory calibrations was limited (two), a Type B evaluation was used to estimate the uncertainty $u_{lts}(a_{ij})$ for each gauge.

At a given target pressure, the variation due to long-term shifts was modeled by a normal distribution such that the best estimated value is $((a_{iPL})_{\max} + (a_{iPL})_{\min})/2$ and there is a 2 out of 3 chance the calibration ratio lies in the interval between maximum and minimum values of a_{iPL} obtained from two calibrations at the pilot laboratory. Then the standard uncertainty due to this source of error equals one-half the difference between the maximum and minimum values:

$$u_{lts}(a_{ij}) = \left((a_{iPL})_{\max} - (a_{iPL})_{\min}\right) \big/ 2 \tag{23}$$

This estimate is unaffected by any systematic bias in the pilot laboratory primary standard, which would be present in all three calibrations. However the estimate does assume that the observed shifts in the calibration ratios are primarily due to the gauges and not the primary standard at the pilot laboratory.

In order to check the latter assumption, results shown in Figures 2 and 3 for the pilot laboratory calibrations are combined in Figure 6 where calibration ratios for gauge 2 are plotted against those for gauge 1. The results correspond to calibrations of CDGs (upper graph) at target pressures of 1 Pa, 3 Pa, 7 Pa, 10 Pa, 30pa, 70 Pa, and 100 Pa, and of RSGs (lower graph) at target pressures of 100 Pa, 300 Pa, 700 Pa and 1000 Pa. Individual points for a given calibration represent results for a pair of gauges at different target pressures where the black symbols refer to Package A and the gray symbols refer to Package B. Dashed lines connect the results for two calibrations of CDGs at 10 Pa and RSGs at 300 Pa in Package A (white + symbols) and Package B (black + symbols). A correlation is seen between CDG2 and CDG1 for a given calibration but this is due to rather similar non-linear behavior of the gauges as a function of pressure. If the observed shifts were due to a common source, such as the pilot laboratory primary standard, they would have the same sign and magnitude for each pair of gauges at a given target pressure (all dashed lines would be aligned diagonally). As may be seen, there is essentially no evidence of a correlation between calibrations at a given target pressure for the CDGs and little evidence of correlation between the RSGs, which are significantly more stable. As shown later in this section (see Figure 7) the long-term instabilities of the RSGs and the re-scaled CDGs are of the same order as the uncertainty due to systematic effects in the primary standard at the pilot laboratory. Several UIMs developed at the pilot laboratory have been checked by direct comparison and their stability found to be consistent with their stated uncertainties. Therefore in the absence of definitive evidence to the contrary, the statement that the observed shifts in the calibration ratios are primarily due to the gauges is taken to be a valid assumption for this key comparison.

The possibility exists that long-term shifts associated with Equation (23) and short-term random effects associated with Equation (22) refer to the same source of variability. If true, the uncertainty contributions from short-term and long-term variability in a_{ij} would not be independent and should be included only once in estimating the combined uncertainty in the normalized gauge readings. As a check, an ANOVA was performed on the 10 mean ratios obtained from two calibrations at the pilot laboratory for each gauge and at each target pressure to test the hypothesis that the "true" means of the two groups of five ratios are equal. For nearly all pilot laboratory data on CDGs, the hypotheses were rejected with a less than 5 % chance that the means are actually equal. For RSGs the hypotheses were rejected with a 40 % chance on average that the means are equal. Although there was a significant chance of equal means for the RSGs (due to their superior calibration stability) it is important that an underestimation of uncertainty in the results be avoided. Therefore contributions from both long-term shifts and short-term random errors in the a_{ij}, for the RSGs as well as the CDGs, were included in estimates of combined uncertainty in the normalized pressure readings of the gauges. The variability due to long-term shifts in gauge response was assumed to be random and uncorrelated for each pair of gauges.

The normalized calibration ratio, $a'_{ij} = n_{iA}\,a_{ij} = n_{iA}(a_{iPL2}, b_{iPL2}, a_{iPL3}, b_{iPL3})\,a_{ij}$, is not only a function of the calibration ratio determined by primary standard j but, through the mean normalization coefficient[11], is also a function of the calibration ratios a_{iPL} and b_{iPL} obtained from two simultaneous calibrations at the pilot laboratory (NIST #2 and #3). Taking into account the propagation of uncertainties, the uncertainty in the normalized calibration coefficients for Package A due to long-term shifts may be estimated from

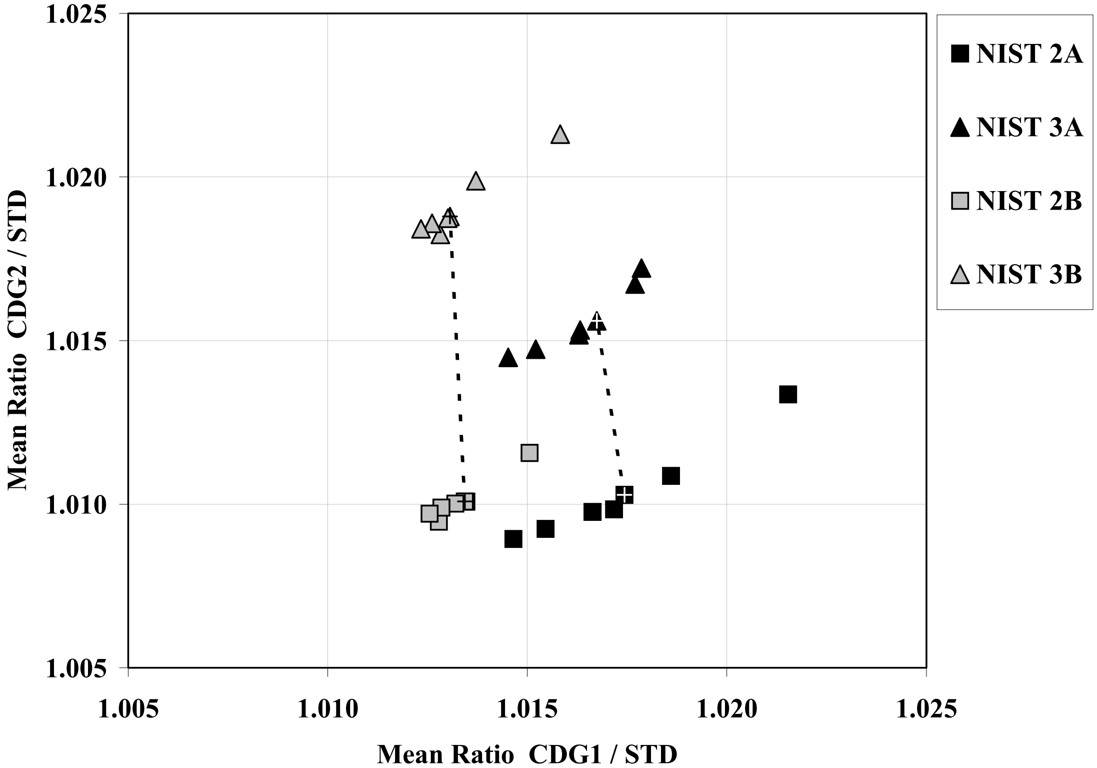
$$u_{lts}^2(a'_{ij}) \simeq n_{iA}^2\,u_{lts}^2(a_{ij}) + (1/8)\,u_{lts}^2(a_{ij}) + (1/8)\,u_{lts}^2(b_{ij}) - r_{jPL}\,(n_{iA}/2)\,u_{lts}^2(a_{ij}) \qquad (24)$$

where $u_{lts}(a_{iPL2}) = u_{lts}(a_{iPL3}) = u_{lts}(a_{ij})$ and $u_{lts}(b_{iPL2}) = u_{lts}(b_{iPL3}) = u_{lts}(b_{ij})$. The partial derivatives (e.g., $\partial a'_{ij}/\partial a_{iPL2}$, etc.) were evaluated using the following approximations: $a_{iPL2} \simeq a_{iPL3} \simeq a_{ij}$, $b_{iPL2} \simeq b_{iPL3} \simeq b_{ij}$, and $(b_{iPL}/a_{iPL}) \approx 1$. The correlation coefficient r_{jPL} is equal to unity when j is either $PL2$ or $PL3$ (NIST #2 or NIST #3) but is zero for all other calibrations of Package A. Similarly for Package B:

$$u_{lts}^2(b'_{ij}) \simeq n_{iB}^2\,u_{lts}^2(b_{ij}) + (1/8)\,u_{lts}^2(b_{ij}) + (1/8)\,u_{lts}^2(a_{ij}) - r_{jPL}\,(n_{iB}/2)\,u_{lts}^2(b_{ij}) \qquad (25)$$

---

[11] The mean is derived from Equation (14) as $n_{iA} = (1/2)\left[\left(\dfrac{a_{iPL2}+b_{iPL2}}{2a_{iPL2}}\right)+\left(\dfrac{a_{iPL3}+b_{iPL3}}{2a_{iPL3}}\right)\right] = (1/4)\left[2+\left(\dfrac{b_{iPL2}}{a_{PL2}}\right)+\left(\dfrac{b_{iPL3}}{a_{PL3}}\right)\right]$.



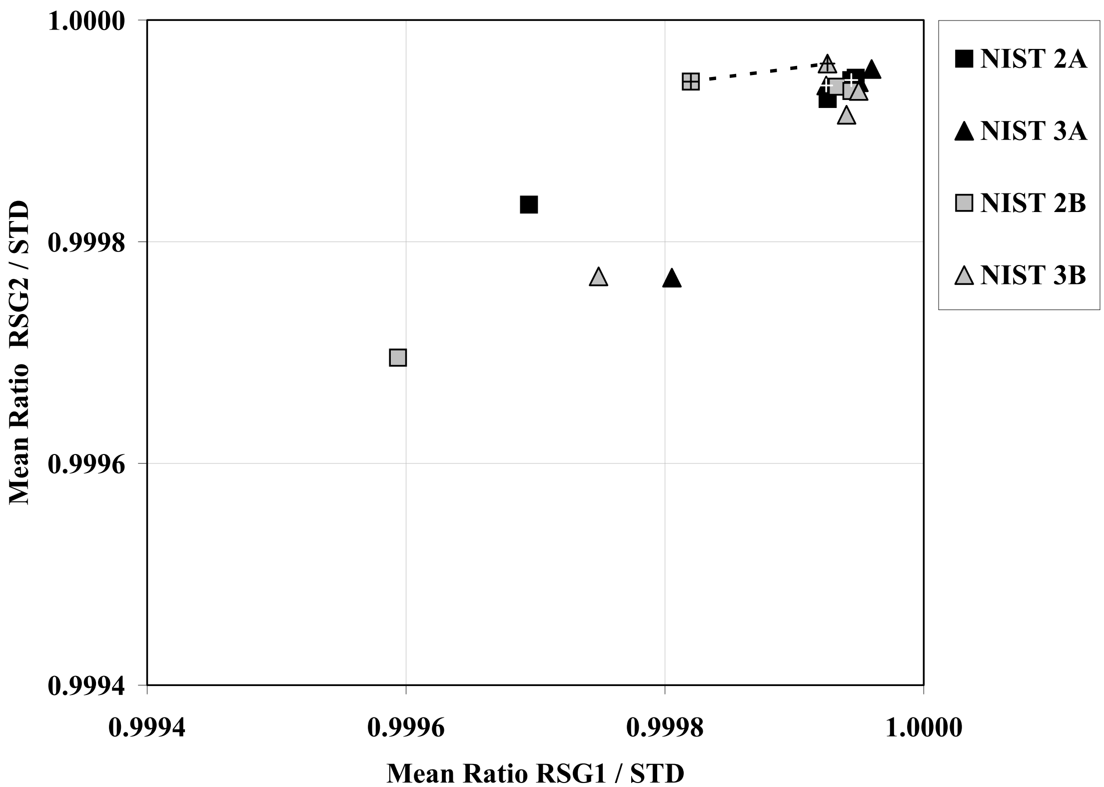


Figure 6. Calibration ratios from pilot laboratory data for gauge 2 versus those for gauge 1. Individual points refer to data at different target pressures. Dashed lines connect results for CDGs at 10 Pa and for RSGs at 300 Pa.

Table 4. Relative standard uncertainty in the calibration ratio a_{ij} or b_{ij} and in the normalized calibration ratio a'_{ij} or b'_{ij} due to long-term shifts in gauge response. The values in parentheses refer to the uncertainty in the normalized ratios obtained from either of the two simultaneous calibrations at the pilot laboratory, NIST #2 or NIST #3. Values above the dotted line refer to CDGs and those below refer to RSGs. Not all digits are significant but are retained for calculation of final results.

Target Press. Pa	$100\times u_{lts}(a_{ij})/a_{ij}$		$100\times u_{lts}(b_{ij})/b_{ij}$		$100\times u_{lts}(a'_{ij})/a'_{ij}$				$100\times u_{lts}(b'_{ij})/b'_{ij}$			
	CDG1 / RSG1	CDG2 / RSG2	CDG1 / RSG1	CDG2 / RSG2	CDG1 / RSG1		CDG2 / RSG2		CDG1 / RSG1		CDG2 / RSG2	
1	0.1682	0.0876	0.0558	0.0538	0.1795	(0.1344)	0.0948	(0.0718)	0.0840	(0.0742)	0.0649	(0.0526)
3	0.0325	0.0119	0.0298	0.0576	0.0360	(0.0277)	0.0240	(0.0224)	0.0336	(0.0262)	0.0612	(0.0457)
10	0.0224	0.0152	0.0003	0.0038	0.0237	(0.0177)	0.0162	(0.0121)	0.0079	(0.0079)	0.0067	(0.0062)
30	0.0033	0.0029	0.0087	0.0045	0.0047	(0.0041)	0.0034	(0.0028)	0.0093	(0.0070)	0.0048	(0.0037)
100	0.0055	0.0033	0.0078	0.0037	0.0065	(0.0052)	0.0037	(0.0029)	0.0085	(0.0064)	0.0041	(0.0031)
300	0.0010	0.0002	0.0053	0.0008	0.0021	(0.0020)	0.00038	(0.00034)	0.0056	(0.0042)	0.00087	(0.00065)
1000	0.00012	0.00021	0.00028	0.00021	0.00016	(0.00014)	0.00022	(0.00016)	0.00030	(0.00023)	0.00023	(0.00018)

The relative uncertainties in calibration ratios due to long-term shifts in gauge response were estimated using Equations (23) to (25) and are given in Table 4. The estimates for CDGs are based on variability of their calibration ratios after re-scaling to the RSGs.

It is noteworthy that the relative uncertainty in the normalized gauge readings [equal to the relative uncertainty in the corresponding normalized ratio via Equation (20)] is of the same order of magnitude as the relative uncertainty due to systematic effects in the primary standards. This is illustrated in Figure 7 where the relative uncertainties of the transfer standard gauge readings are superimposed upon the relative uncertainties of the primary standards (shown in Figure 5). This plot shows that the long-term stability of the transfer standard over the course of this comparison should be sufficient to resolve any relative biases between different primary standards.

Finally, the combined uncertainty in the normalized gauge readings, p_{ij}, at each target pressure was estimated by using data from Tables 3, 4, and 5 and Equations (18) to (20), and is given in Table 5.

6.6. ESTIMATES OF UNCERTAINTIES IN THE CORRECTED MEAN GAUGE READINGS

The component uncertainties in $u_c(p_{ij})$ will also propagate to the combined uncertainty in the corrected mean gauge reading p_j calculated via Equation (17). For the non-pilot laboratories, the combined uncertainty was estimated from [10,11]

$$u_c^2(p_j) \simeq u_c^2(p_{jU}) = u_{std}^2(p_j) + \sum_{i=1}^{2} c^2 u_{rdm}^2(p_{ij}) + \sum_{i=1}^{2} c^2 u_{lts}^2(p_{ij}) \tag{26}$$

where $u_{std}(p_{1j}) = u_{std}(p_{2j}) = u_{std}(p_{jU}) = u_{std}(p_j)$, $c = ½$ is the (common) value for the partial derivatives, $\partial p_{jU}/\partial p_{ij}$, and the approximation $f_C \simeq 1$ was used.

For the pilot laboratory, p_{jU} is the mean of eight values of p_{ij}^{mn} at target pressures up to and including 1000 Pa, where m is the package label and n is the calibration number [see discussion following Equation (16)]. In this case the combined uncertainty in p_j was estimated from:

$$u_c^2(p_j) \simeq u_c^2(p_{jU}) = u_{std}^2(p_j) + \sum_{m=A}^{B}\sum_{n=1}^{2}\sum_{i=1}^{2} c^2 u_{rdm}^2(p_{ij}^{mn}) + \sum_{m=A}^{B}\sum_{n=1}^{2}\sum_{i=1}^{2} c^2 u_{lts}^2(p_{ij}^{mn}) \tag{27}$$

where $c = 1/8$. Note that multiple calibrations at the pilot laboratory tend to reduce the influence of the uncorrelated uncertainties arising from short-term and long-term variability of the gauges on the combined uncertainty in p_j for the pilot laboratory.

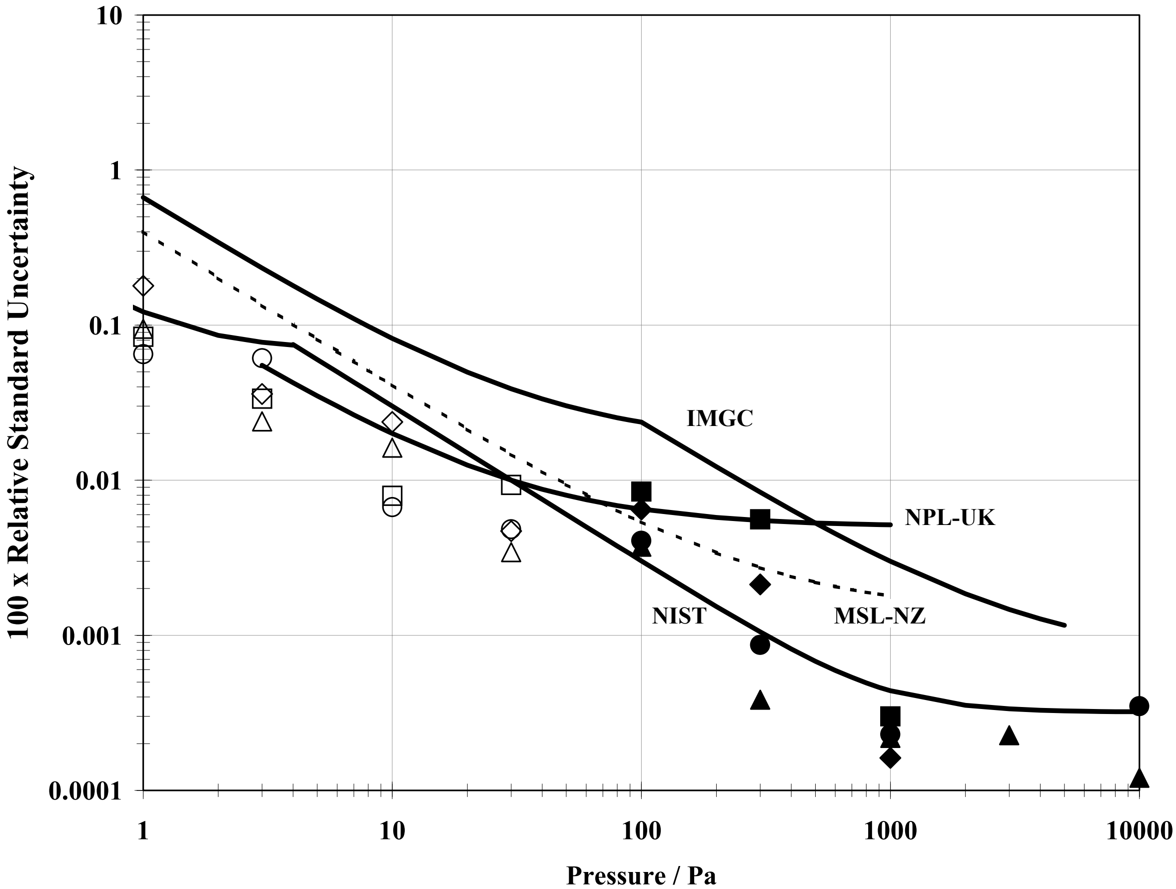


Figure 7. Comparison of the relative uncertainty due to long-term shifts in the transfer standard gauges with the relative uncertainties due to systematic effects in the primary standards. The solid symbols refer to RSGs, the open symbols to (re-scaled) CDGs. Diamond and triangle symbols refer to gauges 1 and 2 in Package A. Square and circular symbols refer to gauges 1 and 2 in Package B.

6.7. EVALUATION OF DEGREES OF EQUIVALENCE

The Mutual Recognition Arrangement (MRA) [1] proposes that the degree of equivalence of a national measurement standard may be stated in two ways, degree of equivalence relative to a key comparison reference value (KCRV) and degree of equivalence between pairs of national standards. Several procedures can be used to define a KCRV each having both advantages and disadvantages, as described in the Appendix. The definition of a KCRV at each target pressure is given by Equation (A1) of the Appendix, which in effect sets the reference value numerically equal to the target pressure.

The degree of equivalence of primary standard j relative to a KCRV is expressed at each target pressure by two quantities, the deviation of p_j from the reference value p_R

$$D_j = p_j - p_R \tag{28}$$

and the expanded uncertainty of this deviation, which is estimated from

$$U_j^2 = k_{95}^2 u_c^2(D_j) = k_{95}^2\left[\left(1-1/N\right)u_c^2(p_j) + u_c^2(p_R)\right] \tag{29}$$

where $u_c(D_j)$ is the combined standard uncertainty of this deviation, k_{95} is the coverage factor that approximates a 95 % level of confidence for the interval defined by U_j, $u_c(p_j)$ and $u_c(p_R)$ are the combined uncertainties in the corrected mean gauge readings and the reference value given by Equations

(27) or (28) and (A5), respectively, and N is the number of primary standards of a given type and is equal to either N_{MAN} or N_{DPB} depending on whether primary standard j is a manometer or a double pressure balance. The term involving $-1/N$ is a correction for the correlation between p_R and p_j.

Following the wording given in the MRA, the degree of equivalence between pairs of primary standards j and j' may be expressed at each target pressure by two quantities, the difference of their deviations from the reference value[12]

$$D_{jj'} = D_j - D_{j'} = (p_j - p_R) - (p_{j'} - p_R) = p_j - p_{j'} \tag{30}$$

and the expanded uncertainty of this difference, which is estimated from

$$U_{jj'}^2 = k_{95}^2 u_c^2(D_{jj'}) = k_{95}^2[u_c^2(p_j) + u_c^2(p_{j'})] \tag{31}$$

where $u_c(D_{jj'})$ is the combined standard uncertainty of this difference, k_{95} is the coverage factor that approximates a 95 % level of confidence for the interval defined by $U_{jj'}$, $u_c(p_j)$ and $u_c(p_{j'})$ are the combined uncertainties in the corrected mean gauge readings obtained with primary standards j and j', respectively, and are estimated from Equation (26) or from Equations (26) and (27).

Values for coverage factors k_{95} that produce the expanded uncertainties U_j and $U_{jj'}$ were obtained using a conventional procedure described in Section A2 of the Appendix and are given in Table A2.

7. RESULTS FOR KEY COMPARISON CCM.P-K5

7.1. COMPARISON OF NORMALIZED GAUGE READINGS

Table 5 presents a summary of the normalized gauge readings, p_{ij}, obtained from calibrations at the pilot and other participant laboratories as a function of nominal target pressures. Results obtained with Package A or with Package B are delineated in the table by a heavy separator line and are presented in chronological order of the calibrations. The calibration ratios for the CDGs, before and after re-scaling to the RSGs, were calculated using Equations (4) and (7), respectively, and are given in columns three/four and five/six. The ratios for the RSGs calculated via Equation (4) are also given in columns five/six. Uncertainties in the ratios due to short-term random effects, which were obtained by means of Equation (22), are given in columns seven/eight. Values of p_{ij}, which were calculated via Equation (15), are given in columns nine/ten. The combined standard uncertainties, $u_c(p_{ij})$, which were calculated according to Equation (18), are given in columns eleven/twelve.

The results for the normalized gauge readings, p_{ij}, and their standard ($k = 1$) uncertainties, $u_c(p_{ij})$, are presented in Figures 8 through 11 in the form of Youden plots [12] in which the difference $p_{2j} - p_t$ is plotted as a function of $p_{1j} - p_t$. The y- and x-axes are labeled as *CDG2 - STD* and *CDG1 - STD* or as *RSG2 - STD* and *RSG1 - STD* for greater clarity. Residual errors associated with normalizing gauge readings from different packages to a common reading manifest themselves as differences between the normalized results obtained from two simultaneous calibrations of the two packages at the pilot laboratory. These differences, although small, can be seen in Figures 8 through 11 as differences between NIST 2A and NIST 2B results (black and gray square symbols) and between NIST 3A and NIST 3B results (black and gray triangle symbols). At several target pressures (e.g., 1 Pa, 100 Pa, and 300 Pa), there is a sizeable scatter of results along the direction of the "precision" diagonal[13] of the plots. This is probably due to the sequential reading (~10 to 15 s time interval) of individual gauges of each pair, (e.g., CDG1 and CDG2) in the presence of temperature-induced pressure fluctuations at nominal line pressures of 100 kPa.

---

[12] The degree of equivalence between pairs of standards is written as stated in the MRA but in reality the difference $D_{jj'}$ does not require the calculation of a KCRV.

[13] Defined in Section 8.

Table 5. Summary of key comparison results for calibration ratios, $a_{i\,j}$ and $b_{i\,j}$, their uncertainty due to short-term random effects, $u_{rdm}(a_{i\,j})$ and $u_{rdm}(b_{i\,j})$, calculated values for normalized readings of gauge i, $p_{i\,j}$, when the pressure measured/generated by primary standard j equals the target pressure, and their combined standard uncertainty, $u_c(p_{i\,j})$. Values above the dotted line refer to CDGs and those below refer to RSGs. Not all digits are significant but are retained for calculation of final results.

NMI	Target Press. Pa	Calibration Ratios Before Re-scaling CDG1	Calibration Ratios Before Re-scaling CDG2	$a_{i\,j}$ or $b_{i\,j}$ CDG1 / RSG1	$a_{i\,j}$ or $b_{i\,j}$ CDG2 / RSG2	u_{rdm} ($a_{i\,j}$ or $b_{i\,j}$) CDG1 / RSG1	u_{rdm} ($a_{i\,j}$ or $b_{i\,j}$) CDG2 / RSG2	$p_{i\,j}$ / Pa CDG1 / RSG1	$p_{i\,j}$ / Pa CDG2 / RSG2	u_c ($p_{i\,j}$) / Pa CDG1 / RSG1	u_c ($p_{i\,j}$) / Pa CDG2 / RSG2
NIST #2A	1	1.02153	1.01336	1.00648	1.00421	0.00097	0.00083	1.0054	1.0036	0.0021	0.0016
	3	1.01860	1.01086	1.00359	1.00174	0.00042	0.00038	3.0071	3.0034	0.0028	0.0027
	10	1.01744	1.01029	1.00244	1.00117	0.00016	0.00016	10.0156	10.0071	0.0038	0.0036
	30	1.01663	1.00976	1.00165	1.00065	0.00008	0.00010	30.0300	30.0108	0.0040	0.0042
	100	1.01465	1.00894	0.999695	0.999834	0.000035	0.000071	99.9655	99.9799	0.0069	0.0083
	300			0.999944	0.999946	0.000016	0.000024	299.9740	299.9851	0.0084	0.0080
	1000			0.999947	0.999948	0.000004	0.000008	999.9466	999.9431	0.0062	0.0094
NPL-UK	3	1.01775	1.01106	1.00317	1.00211	0.00045	0.00054	3.0058	3.0045	0.0024	0.0024
	10	1.01668	1.01002	1.00212	1.00108	0.00016	0.00026	10.0125	10.0062	0.0035	0.0037
	30	1.01618	1.00973	1.00163	1.00079	0.00011	0.00021	30.0293	30.0150	0.0046	0.0070
	100	1.01387	1.00859	0.999349	0.999656	0.000080	0.000190	99.9310	99.9622	0.0122	0.0204
	300			0.999783	0.999831	0.000037	0.000056	299.9256	299.9507	0.0209	0.0236
	1000			0.999906	0.999945	0.000010	0.000016	999.9052	999.9396	0.0525	0.0540
IMGC	1	1.01944	1.02739	1.00645	1.01622	0.00845	0.00902	1.0054	1.0156	0.0109	0.0111
	3	1.01706	1.01474	1.00410	1.00371	0.00242	0.00285	3.0086	3.0093	0.0101	0.0110
	10	1.01579	1.01323	1.00284	1.00221	0.00078	0.00071	10.0196	10.0175	0.0116	0.0110
	30	1.01508	1.01213	1.00215	1.00113	0.00028	0.00030	30.0449	30.0251	0.0144	0.0147
	100	1.01292	1.01106	1.000011	1.000064	0.000084	0.000151	99.9972	100.0030	0.0260	0.0284
	300			0.999992	1.000035	0.000039	0.000055	299.9883	300.0119	0.0285	0.0301
	1000			0.999958	0.999964	0.000016	0.000024	999.9571	999.9593	0.0342	0.0383
NIST #3A	1	1.01786	1.01721	1.00309	1.00245	0.00090	0.00099	1.0020	1.0018	0.0020	0.0017
	3	1.01770	1.01673	1.00294	1.00198	0.00035	0.00024	3.0051	3.0041	0.0027	0.0025
	10	1.01674	1.01560	1.00199	1.00086	0.00013	0.00010	10.0112	10.0040	0.0037	0.0034
	30	1.01633	1.01532	1.00158	1.00059	0.00008	0.00010	30.0280	30.0091	0.0041	0.0043
	100	1.01452	1.01449	0.999805	0.999768	0.000072	0.000090	99.9766	99.9733	0.0093	0.0099
	300			0.999925	0.999941	0.000021	0.000027	299.9682	299.9837	0.0092	0.0087
	1000			0.999950	0.999944	0.000005	0.000008	999.9490	999.9390	0.0070	0.0097
NIST #2B	1	1.01506	1.01157	1.00207	1.00153	0.00047	0.00073	1.0031	1.0022	0.0015	0.0015
	3	1.01347	1.01008	1.00051	1.00006	0.00024	0.00030	3.0052	3.0020	0.0026	0.0028
	10	1.01343	1.01008	1.00047	1.00006	0.00016	0.00010	10.0134	10.0052	0.0035	0.0032
	30	1.01320	1.01002	1.00023	1.00000	0.00015	0.00006	30.0264	30.0086	0.0059	0.0036
	100	1.01255	1.00971	0.999594	0.999695	0.000150	0.000041	99.9633	99.9730	0.0166	0.0060
	300			0.999820	0.999944	0.000052	0.000007	299.9553	299.9820	0.0203	0.0042
	1000			0.999944	0.999936	0.000011	0.000006	999.9450	999.9410	0.0119	0.0076

NMI	Target Press. Pa	Calibration Ratios Before Re-scaling		a_{ij} or b_{ij}		u_{rdm} (a_{ij} or b_{ij})		p_{ij} / Pa		$u_c(p_{ij})$ / Pa	
		CDG1	CDG2	CDG1 / RSG1	CDG2 / RSG2	CDG1 / RSG1	CDG2 / RSG2	CDG1 / RSG1	CDG2 / RSG2	CDG1 / RSG1	CDG2 / RSG2
MSL-NZ	1	1.00779	1.01675	0.99544	0.99976	0.00800	0.00634	0.9965	1.0004	0.0090	0.0075
	3	1.01233	1.01602	0.99992	0.99904	0.00425	0.00271	3.0035	2.9990	0.0134	0.0093
	10	1.01680	1.02062	1.00434	1.00357	0.00155	0.00162	10.0522	10.0403	0.0160	0.0167
	30	1.01251	1.01724	1.00010	1.00024	0.00015	0.00019	30.0224	30.0158	0.0070	0.0075
	100	1.01211	1.01654	0.999701	0.999553	0.000046	0.000059	99.9741	99.9588	0.0110	0.0090
	300			1.000022	0.999910	0.000028	0.000037	300.0158	299.9716	0.0205	0.0140
	1000			0.999957	0.999944	0.000009	0.000009	999.9576	999.9488	0.0203	0.0201
NIST #3B	1	1.01582	1.02131	1.00319	1.00261	0.00051	0.00050	1.0043	1.0032	0.0015	0.0014
	3	1.01371	1.01988	1.00110	1.00121	0.00031	0.00033	3.0070	3.0055	0.0026	0.0029
	10	1.01306	1.01879	1.00046	1.00013	0.00020	0.00015	10.0134	10.0059	0.0037	0.0034
	30	1.01301	1.01874	1.00041	1.00008	0.00013	0.00013	30.0317	30.0113	0.0052	0.0049
	100	1.01234	1.01842	0.999749	0.999769	0.000119	0.000112	99.9788	99.9803	0.0139	0.0120
	300			0.999926	0.999961	0.000031	0.000029	299.9869	299.9868	0.0160	0.0094
	1000			0.999950	0.999936	0.000010	0.000008	999.9506	999.9410	0.0108	0.0092

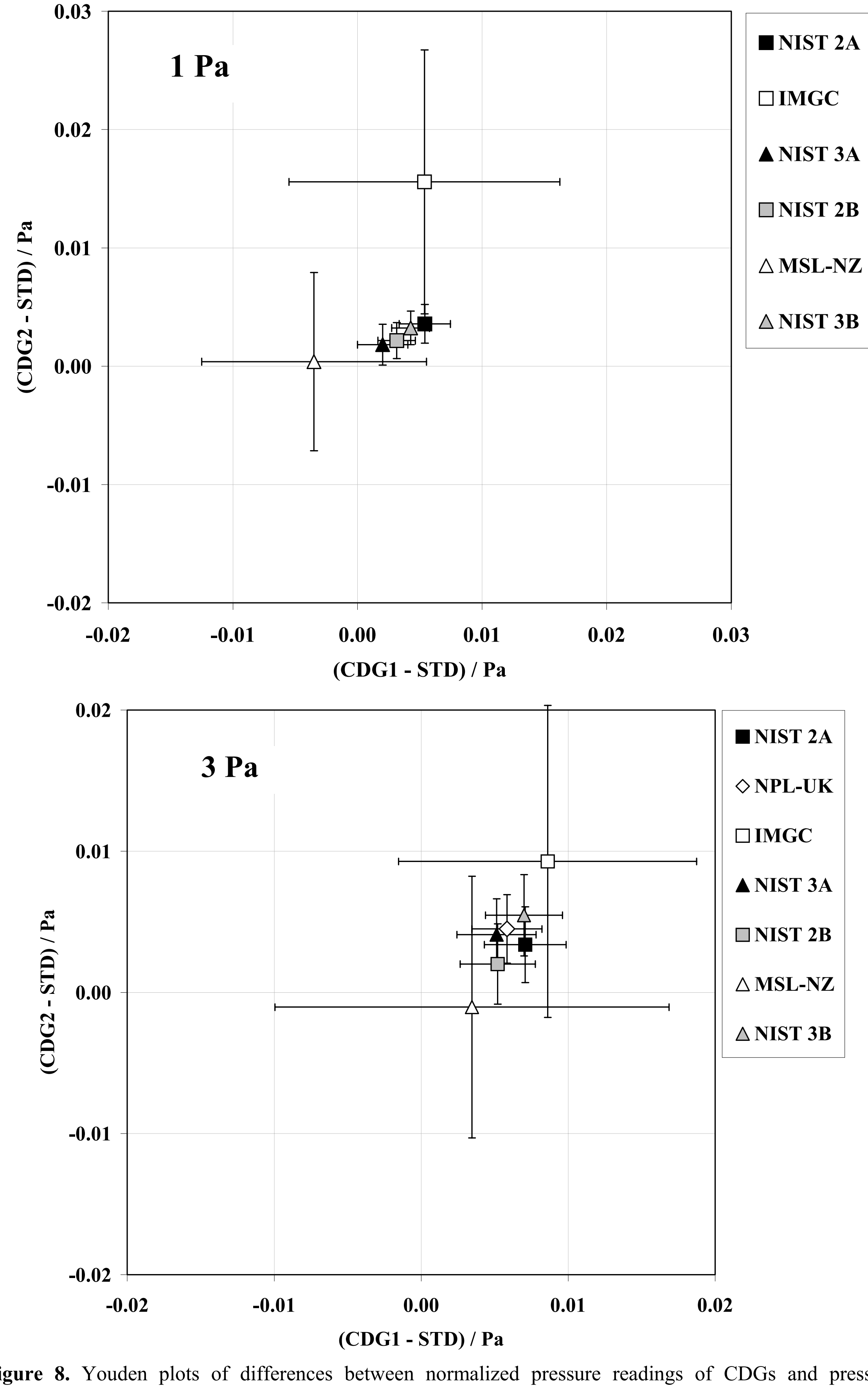


Figure 8. Youden plots of differences between normalized pressure readings of CDGs and pressures measured/generated by primary standards when equal to target pressures of 1 Pa and 3 Pa. The error bars refer to combined standard ($k = 1$) uncertainties.

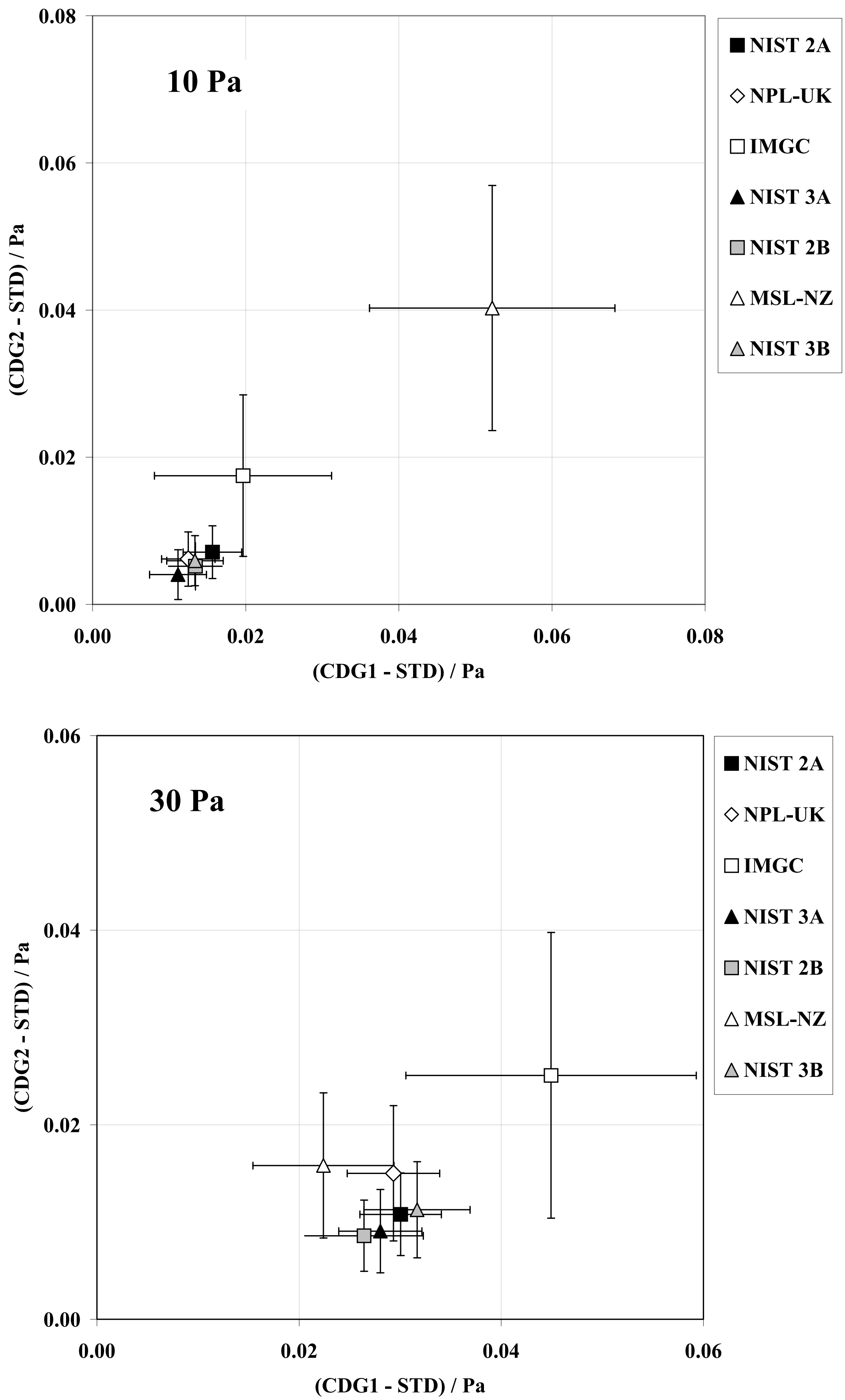
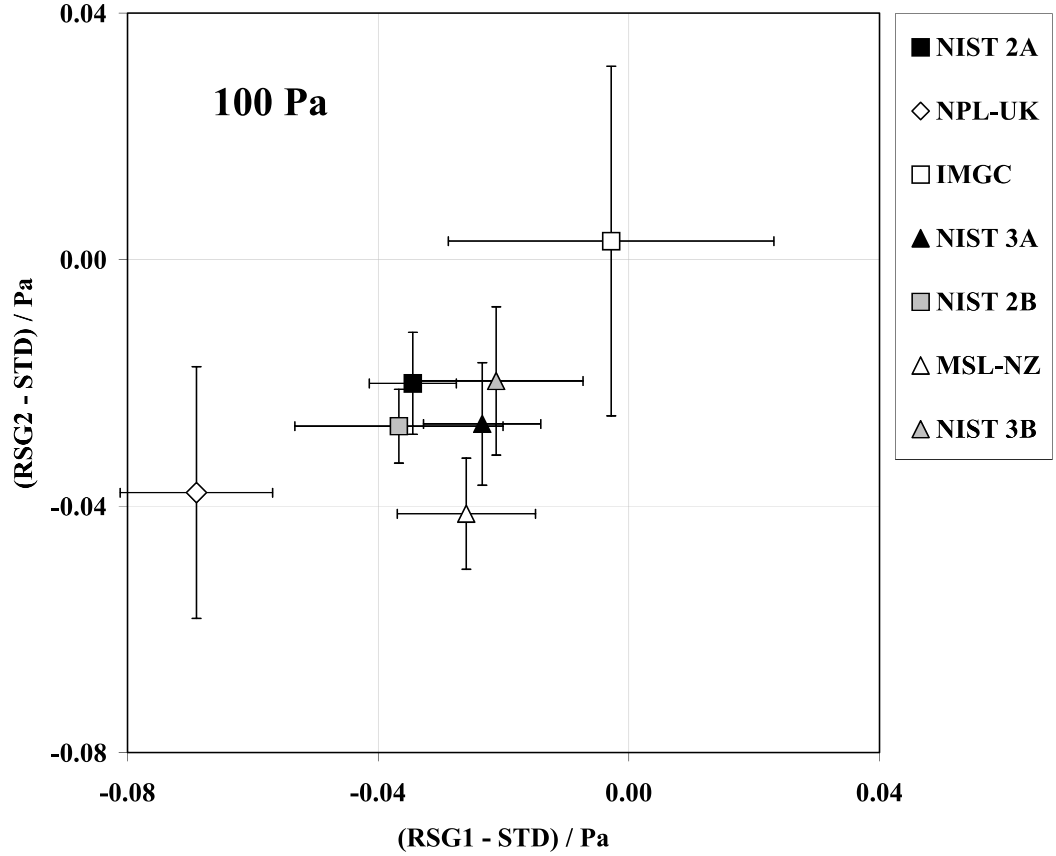


Figure 9. Youden plots of differences between normalized pressure readings of CDGs and pressures measured/generated by primary standards when equal to target pressures of 10 Pa and 30 Pa. The error bars refer to combined standard ($k = 1$) uncertainties.



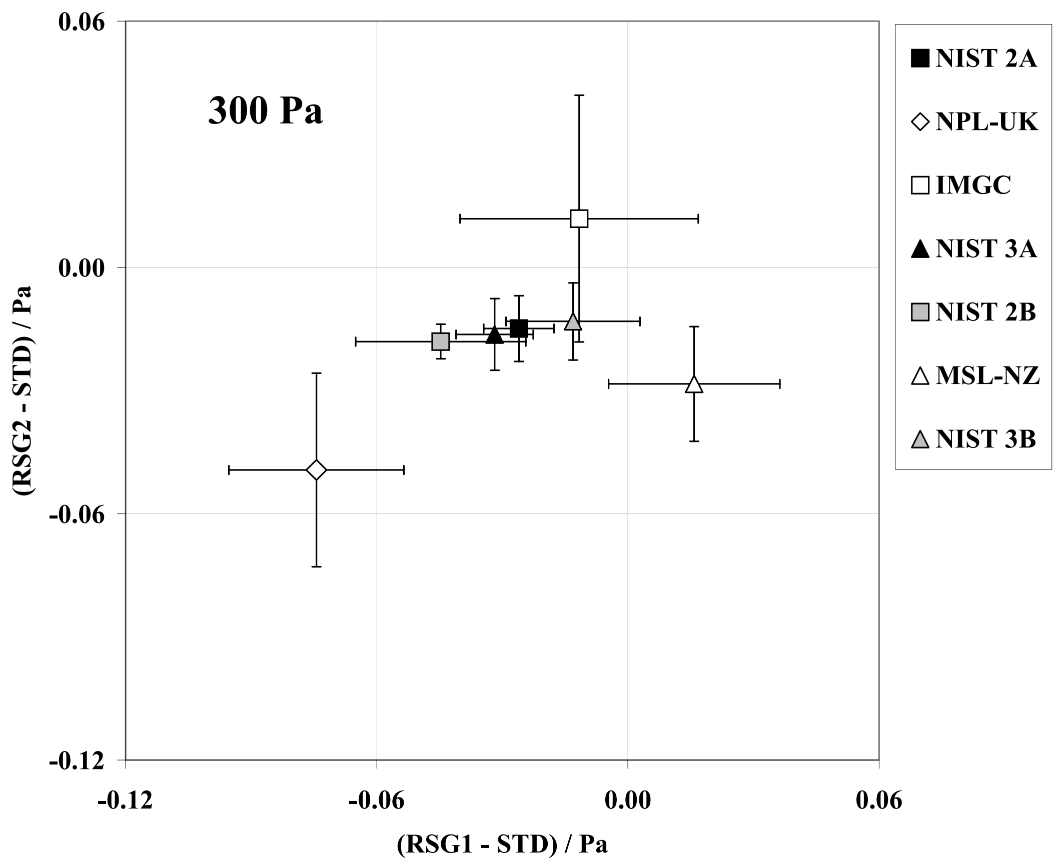


Figure 10. Youden plots of differences between normalized pressure readings of RSGs and pressures measured/generated by primary standards when equal to target pressures of 100 Pa and 300 Pa. The error bars refer to combined standard (k = 1) uncertainties.

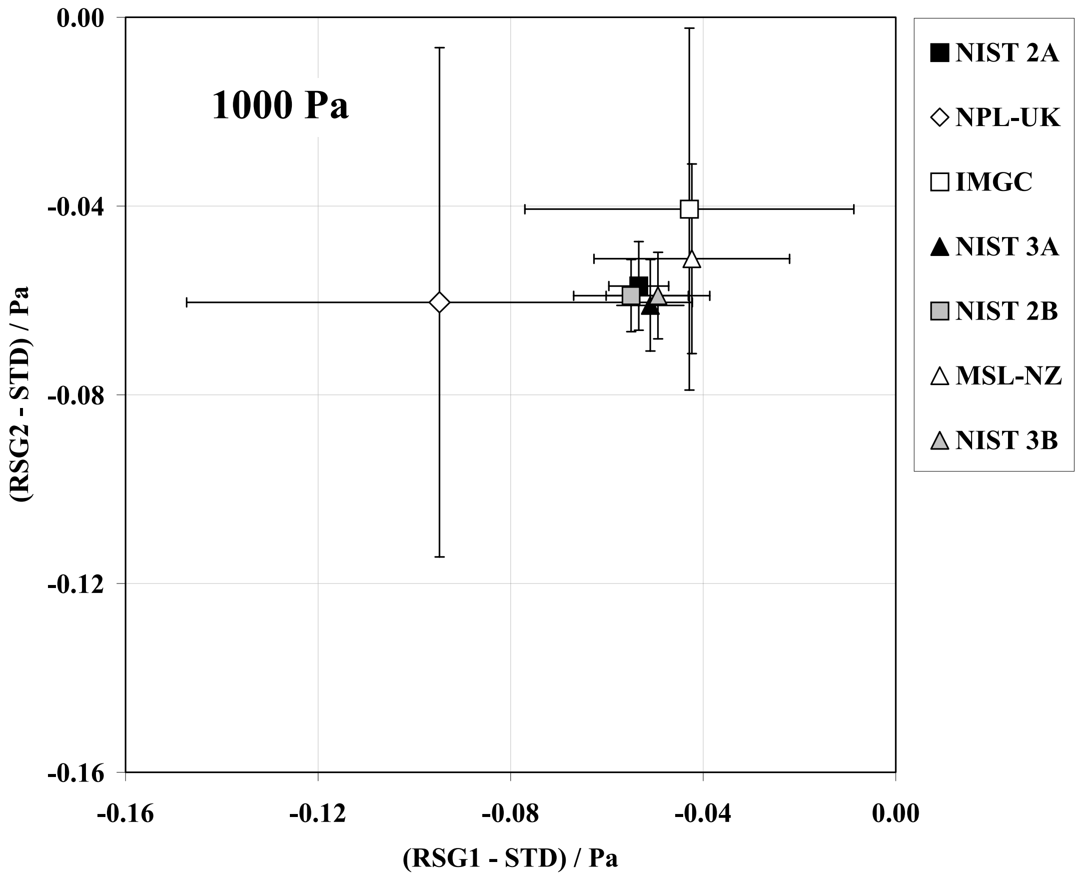


Figure 11. Youden plot of differences between normalized pressure readings of RSGs and pressures measured/generated by primary standards when equal to a target pressure of 1000 Pa. The error bars refer to combined standard (k = 1) uncertainties.

7.2. DEGREES OF EQUIVALENCE OF THE PRIMARY STANDARDS

Table 6 presents a summary of final results for the pilot and participant NMIs as a function of nominal target pressures. The values for the corrected mean gauge readings p_j, which were calculated from Equation (17) using data in Tables 5 and A1, are given in column three. The combined uncertainties $u_c(p_j)$, which were calculated using Equation (26) or (27), are given in column four. The remaining columns present degrees of equivalence of the measurement standards expressed quantitatively in two ways: (1) deviations from reference values, and (2) pairwise differences between these deviations. The deviations D_j were calculated via Equations (28) and (A1) using data from Tables 6 and A1 in which the reference value at 10 Pa does not include the result from MSL-NZ (see Figures 9 and A1, and footnote 18). The expanded uncertainties of these deviations, U_j, were calculated using Equation (29) and data in Tables 6 and A1, and the coverage factors in Table A2. The pairwise differences between the deviations, $D_{jj'}$, and the expanded uncertainties of these differences, $U_{jj'}$, were calculated using Equations (30) and (31) and data from Table 6, and coverage factors from Table A2. The shaded cells in Table 6 indicate pressures at which the condition $|D_j| \leq U_j$ or $|D_{jj'}| \leq U_{jj'}$ is not satisfied.

The degrees of equivalence of individual NMIs with respect to key comparison reference values are presented graphically in Figures 12 to 15 as plots of deviations, $D_j = p_j - p_R$, versus NMI and are summarized in Figure 16 where the ratios, D_j / U_j, for the participating laboratories are plotted as a function of pressure.

Table 6. Degrees of equivalence expressed in two ways: degree of equivalence of an NMI relative to the key comparison reference values, and degree of equivalence between pairs of NMIs. D_j is the deviation of the corrected mean gauge reading p_j obtained by NMI$_j$ from the reference value and U_j is the expanded uncertainty† of this deviation. $D_{j\,j'}$ is the difference between pairs of corrected mean gauge readings from NMI$_j$ and NMI$_{j'}$, and $U_{j\,j'}$ is the expanded uncertainty† of this difference. The shaded cells indicate results for which $|D_j|$ exceeds U_j or $|D_{j\,j'}|$ exceeds $U_{j\,j'}$.

† U_j and $U_{j\,j'}$ refer to a 95% level of confidence

						NMI$_{j'}$							
						IMGC		MSL-NZ		NIST		NPL-UK	
NMI$_j$	**Nominal Pressure Pa**	p_j Pa	$u_c(p_j)$ Pa	D_j Pa	U_j Pa	$D_{j\,j'}$ Pa	$U_{j\,j'}$ Pa	$D_{j\,j'}$ Pa	$U_{j\,j'}$ Pa	$D_{j\,j'}$ Pa	$U_{j\,j'}$ Pa	$D_{j\,j'}$ Pa	$U_{j\,j'}$ Pa
IMGC	1	1.0078	0.0091	0.008	0.015			0.012	0.022	0.007	0.019		
	3	3.0052	0.0090	0.005	0.017			0.008	0.025	0.004	0.019	0.004	0.018
	10	10.0061	0.0099	0.006	0.018			-0.028	0.032	0.009	0.020	0.009	0.020
	30	30.013	0.013	0.013	0.022			0.016	0.028	0.016	0.027	0.013	0.028
	100	100.030	0.026	0.030	0.043			0.034	0.053	0.026	0.051	0.054	0.056
	300	300.017	0.027	0.017	0.047			0.006	0.060	0.022	0.055	0.062	0.066
	1000	1000.011	0.033	0.011	0.060			0.005	0.076	0.014	0.066	0.04	0.12
MSL-NZ	1	0.9958	0.0065	-0.0042	0.0081	-0.012	0.022			-0.005	0.013		
	3	2.9974	0.0086	-0.0026	0.0098	-0.008	0.025			-0.004	0.019	-0.004	0.019
	10	10.034	0.012	0.034	0.026	0.028	0.032			0.037	0.030	0.037	0.027
	30	29.9968	0.0060	-0.0032	0.0078	-0.016	0.028			0.000	0.013	-0.003	0.015
	100	99.9965	0.0081	-0.004	0.012	-0.034	0.053			-0.007	0.018	0.020	0.032
	300	300.011	0.014	0.011	0.018	-0.006	0.060			0.016	0.029	0.056	0.047
	1000	1000.006	0.019	0.006	0.028	-0.005	0.076			0.009	0.039	0.03	0.11
NIST	1	1.0006	0.0013	0.0006	0.0082	-0.007	0.019	0.005	0.013				
	3	3.0011	0.0024	0.001	0.010	-0.004	0.019	0.004	0.019			-0.0002	0.0062
	10	9.9970	0.0031	-0.0030	0.0092	-0.009	0.020	-0.037	0.030			0.0002	0.0083
	30	29.9972	0.0032	-0.0028	0.0092	-0.016	0.027	0.000	0.013			-0.003	0.011
	100	100.0039	0.0048	0.004	0.015	-0.026	0.051	0.007	0.018			0.027	0.029
	300	299.9949	0.0050	-0.005	0.019	-0.022	0.055	-0.016	0.029			0.040	0.040
	1000	999.9970	0.0052	-0.003	0.029	-0.014	0.066	-0.009	0.039			0.02	0.10
NPL-UK	3	3.0014	0.0021	0.001	0.010	-0.004	0.018	0.004	0.019	0.0002	0.0062		
	10	9.9969	0.0029	-0.0031	0.0091	-0.009	0.020	-0.037	0.027	-0.0002	0.0083		
	30	29.9998	0.0047	0.000	0.011	-0.013	0.028	0.003	0.015	0.003	0.011		
	100	99.977	0.013	-0.023	0.025	-0.054	0.056	-0.020	0.032	-0.027	0.029		
	300	299.955	0.020	-0.045	0.036	-0.062	0.066	-0.056	0.047	-0.040	0.040		
	1000	999.975	0.052	-0.025	0.088	-0.04	0.12	-0.03	0.11	-0.02	0.10		

† Value of p_j is not included in the calculation of the reference value and so there is no correlation term $-1/N$ needed in the calculation of U_j.

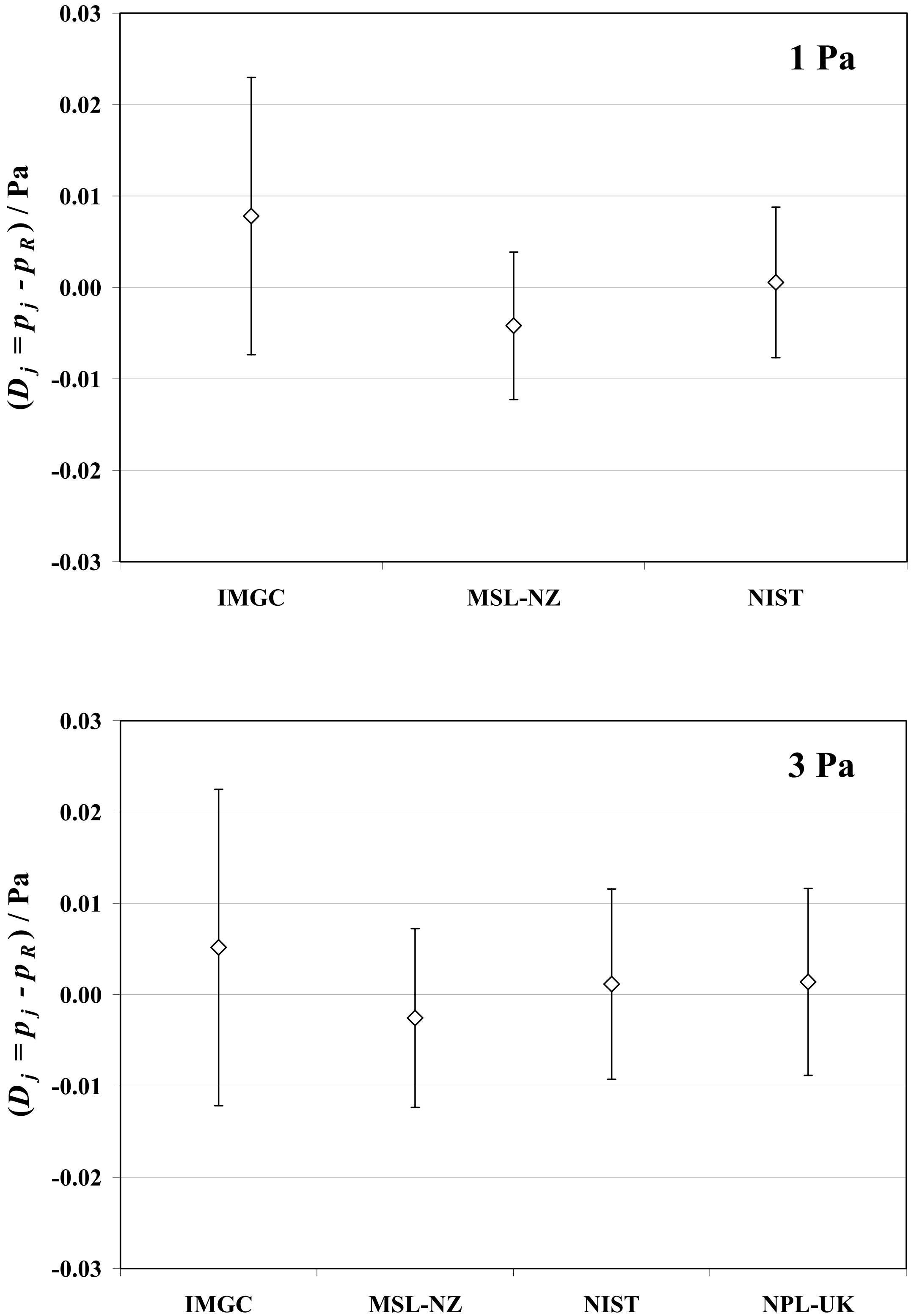
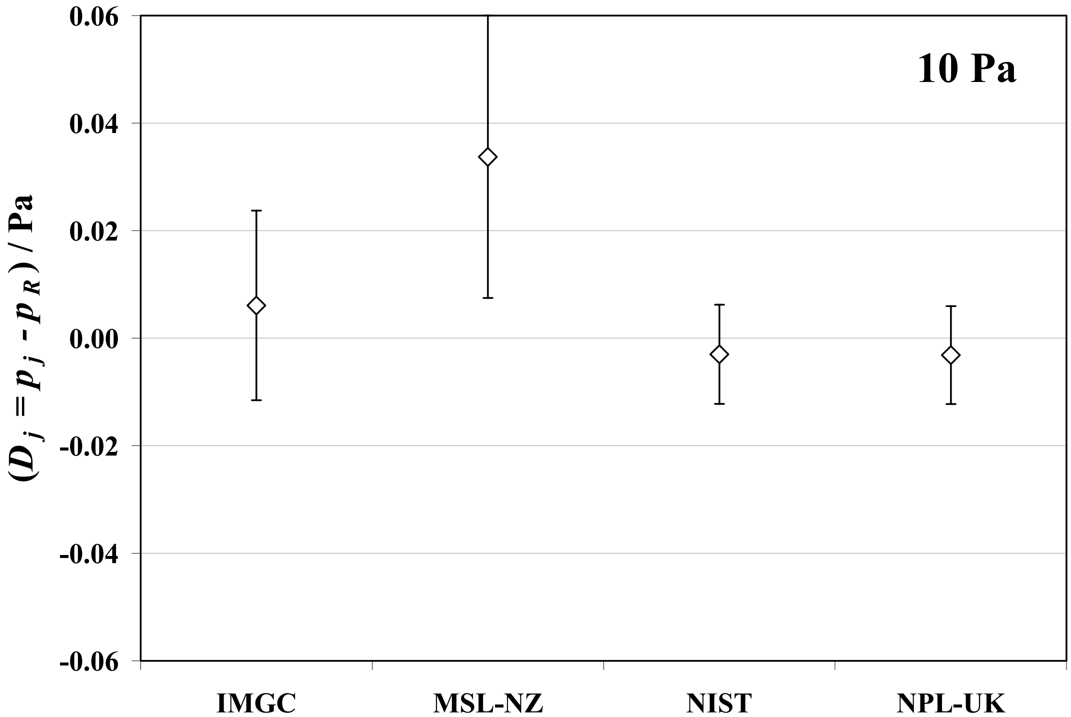


Figure 12. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference values at 1 Pa and 3 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.



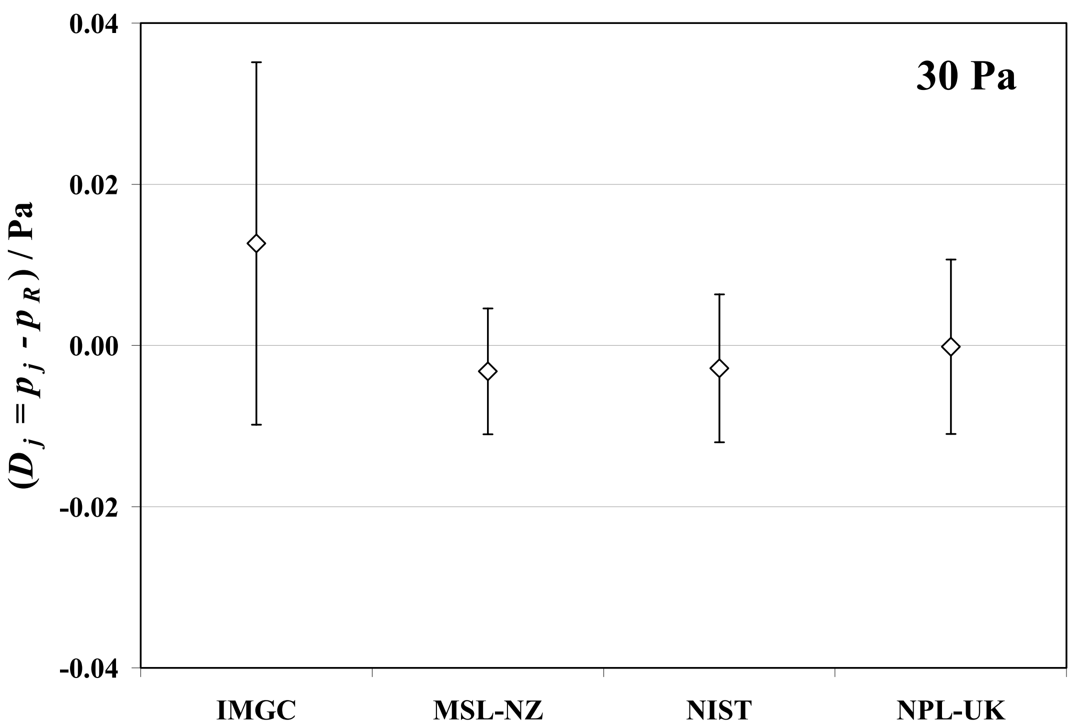
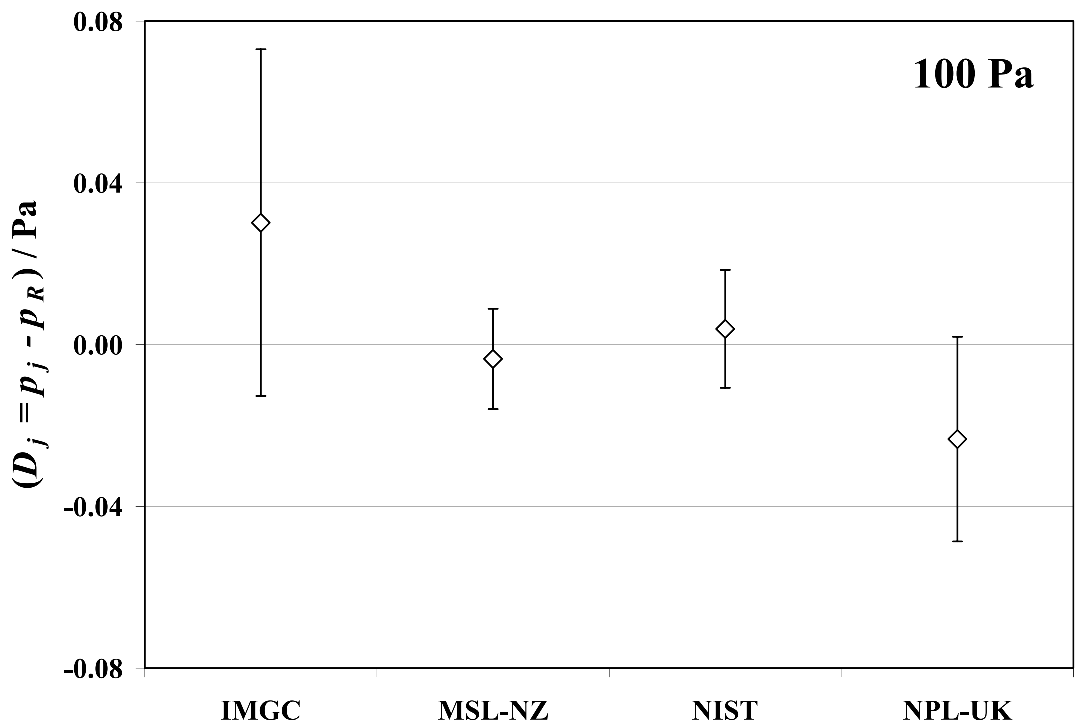


Figure 13. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference values at 10 Pa and 30 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.



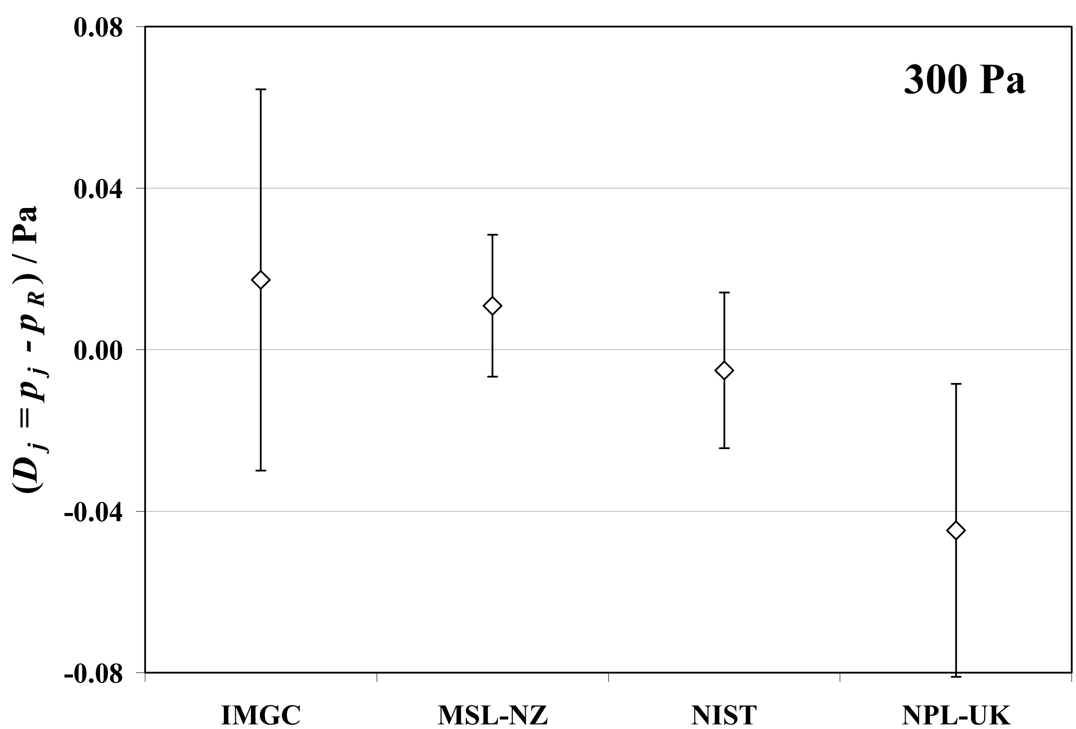


Figure 14. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference values at 100 Pa and 300 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.

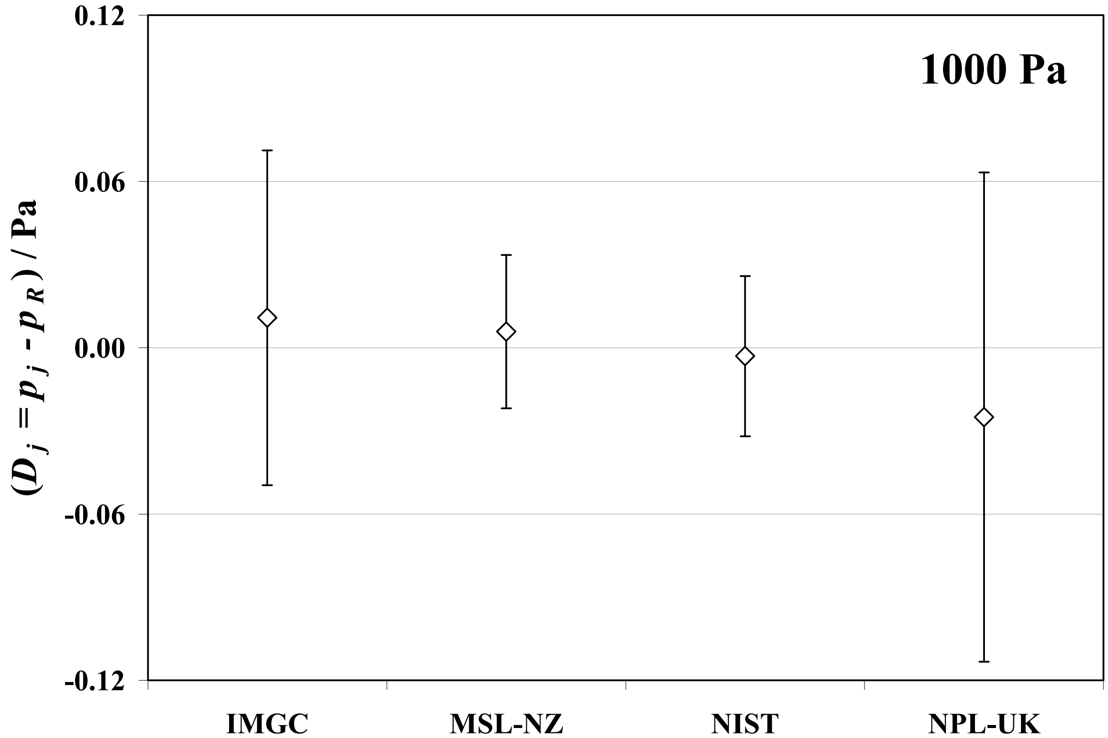


Figure 15. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference value at 1000 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.

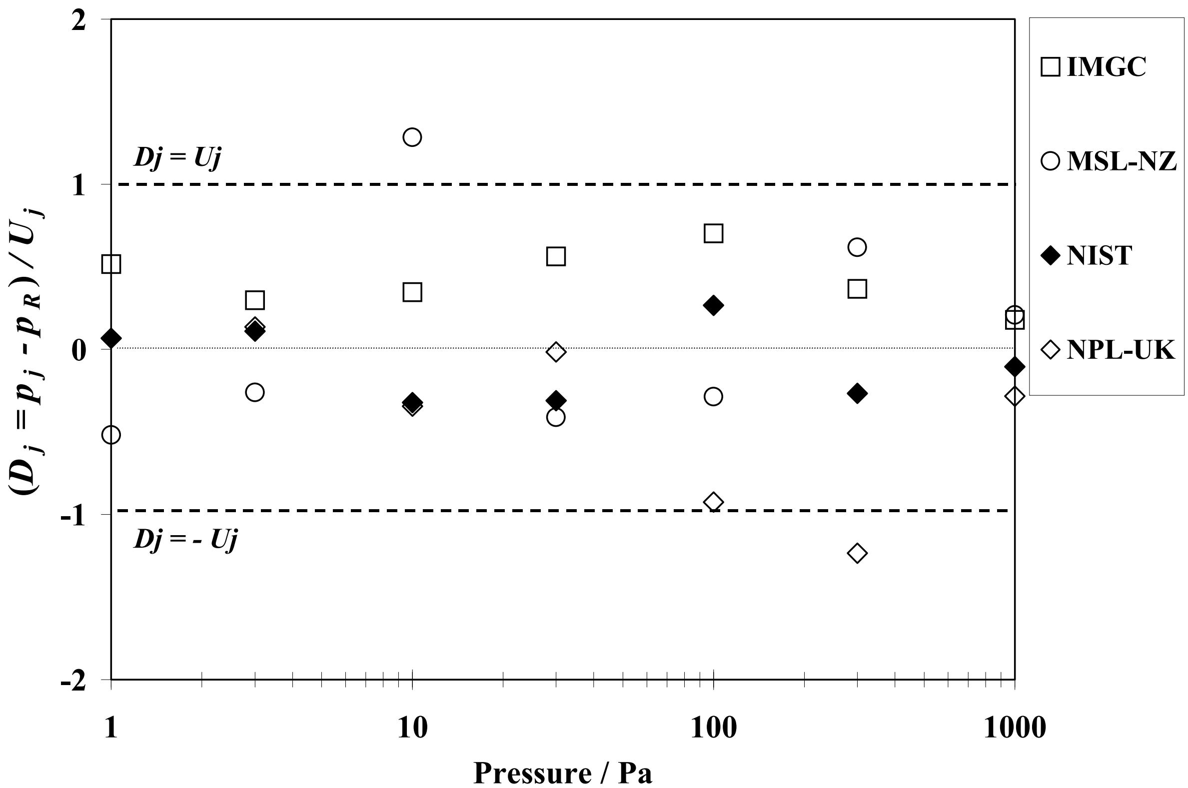


Figure 16. Summary of results for the degrees of equivalence of NMIs with respect to the key comparison reference values. The heavy dashed lines indicate the 95 % confidence interval.

8. DISCUSSION

The use of pairs of pressure transducers in the transfer standard package proved to be valuable not only because of redundancy (e.g., when RSG2 in Package A failed while at IMGC) but more importantly it enabled Youden graphical analyses to be used in the interpretation of the key comparison results. The Youden graphical representation has several important features. If only random errors of precision are present the data points from individual primary standards will be distributed in a circular pattern (in the limit of a large number of standards). However if relative bias between individual primary standards exists, the data points will be distributed along a diagonal at 45 degrees to the positive *y*- and *x*-axes because primary standards that realize "true" pressures that are higher (or lower) will produce higher (or lower) readings in <u>both</u> pressure transducers. The scatter of data in a direction perpendicular to this diagonal provides a measure of precision of the transfer standard gauges. The Youden plots of the present results clearly show that the transfer standard gauges have sufficient precision to differentiate the relative systematic biases between individual primary standards (Figures 8 to 11).

In the present comparison, the degrees of equivalence of the measurement standards were expressed quantitatively in two ways: the deviation of corrected mean gauge readings from a key comparison reference value, D_j, and the pairwise difference between these deviations, $D_{jj'}$. When interpreting the results it is useful to note that $D_{jj'}$ may be regarded as a surrogate for the difference in "true" pressures actually realized by the pair of primary standards when both are set to measure/generate the same target pressure. Similarly, D_j represents the deviation of the "true" pressure realized by primary standard *j* from the corresponding key comparison reference value. However, D_j is not necessarily equal to the deviation of primary standard *j* from the SI value. Although the key comparison reference value is likely to be a close approximation to the SI value, it is possible that some of the results (corrected mean gauge readings) from individual NMIs may be even closer.

In the MRA the term 'degree of equivalence of measurement standards' is taken to mean the degree to which a standard is consistent with a key comparison reference value or with a measurement standard at another laboratory. A measure of the degree of consistency is provided by the relative magnitudes of the deviation D_j and its uncertainty, U_j, or the relative magnitudes of the pairwise difference $D_{jj'}$ and its uncertainty, $U_{jj'}$. The shaded cells in Table 6 indicate pressures at which results do not satisfy the condition $|D_j| \leq U_j$ or $|D_{jj'}| \leq U_{jj'}$. The two cases that do not satisfy the condition with respect to the key comparison reference value, the result from the MSL-NZ at 10 Pa and the result from the NPL-UK at 300 Pa, are marginal and would satisfy the condition if results were rounded to one significant figure.

The results presented in this report are based on data originally submitted to the pilot laboratory for preparation of the Draft A report and as such they represent the operational status of the low differential-pressure standards at the time of the measurements[14]. In hindsight, MSL-NZ found that their data for the 10 Pa point had comparatively large shifts in the generated zero pressures for two of the runs, which may explain the offset in this result.

[14] The Guidelines for CIPM Key Comparisons (Appendix F of the MRA) state that once results have been submitted to the pilot laboratory they stand and can only be changed under unusual circumstances and with the agreement of all participants.

9. CONCLUSIONS

The most critical element in the success of the present comparison of low differential-pressure standards was the use of two different types of transducer as transfer standard artifacts. The combination of resonant silicon gauges with their exceptional calibration stability and capacitance diaphragm gauges to provide pressure resolution yielded transfer standards that had accuracies commensurate with the measurement standards being compared, over the entire pressure range of the comparison. In addition, the gauges were sufficiently rugged to withstand the inevitable rough treatment during shipment between laboratories.

The comparison tested two principal methods used by NMIs to realize their low differential-pressure standards, namely, double pressure balances and liquid-column manometers. The results for one double pressure balance and three manometers revealed no significant relative bias between the principal measurement methods.

Finally, the key comparison established the degrees of equivalence of differential-pressure standards at four NMIs, both with respect to key comparison reference values as well as between pairs of measurement standards. The differential-pressure standards of the participating NMIs were generally found to be equivalent.

APPENDIX

A1. Reference Values for Key Comparison CCM.P-K5

A key comparison reference value (KCRV) may be defined at each target pressure as an average of the mean normalized gauge readings that would be obtained at the different laboratories when their primary standards measure/generate pressures exactly equal to the target value. There are several procedures possible for averaging [13], which include a simple mean of all data, a median of all data and, since the primary standards represent two principal measurement methods, a mean of the measurement method means, or a weighted mean of the measurement method means with weights inversely proportional to estimates of method variance. Each procedure has some advantages and disadvantages.

An arithmetic mean of the combined data has the advantage of simplicity but if the "true" means of the different measurement methods are not the same but have relative bias then the arithmetic mean will weight the methods by "popularity", which is not desirable. Another disadvantage is the simple mean is sensitive to outliers.

A median of all data is relatively insensitive to outliers but it may effectively omit one of the measurement methods from the analysis if there is significant relative bias between methods. The major disadvantage of a median however is the lack of theory on which to base uncertainty estimates.

A major advantage of a mean of the measurement method means is that it incorporates the range of typical values obtained with different measurement methods without weighting by popularity, as does the simple mean of all data. When the population of one of the methods is one as in the present case (one double pressure balance), the lack of a significant difference between method means can still be interpreted as no significant relative bias between the measurement methods. However if a significant difference between the method means is observed the source of bias cannot be uniquely identified[15]. Like the mean of the combined data the mean of the method means is sensitive to the influence of outliers, which can only be eliminated by exclusion from the calculation of the method means.

A weighted mean of the measurement method means, with weights inversely proportional to the "true" method variance, may yield the most precise estimate of the overall mean <u>but</u> the weights must be known without error and any between-method bias must be negligible. Since weights are usually not known without error, using a weighted mean when weights are not known can lead to estimates with less precision than methods based on equal weights. Furthermore, weights (method variance) cannot be estimated when a method is represented by only one primary standard as in the present comparison.

Considering these options, an <u>unweighted</u> mean of the measurement method means was selected as a reasonable procedure to obtain reference values for this key comparison.

As stated earlier (Section 6.4), the "true" pressures realized by the primary standards when set to measure/generate a given target pressure should, on average, closely approximate the SI value under the assumption that deviations from the SI value are randomly distributed. Therefore, it is reasonable to correct the mean normalized gauge readings so that their ensemble average (i.e., the KCRV) is also equal to the target pressure. This correction, in effect, sets the KCRV numerically equal to the target pressure.

At target pressures (p_t) of 1000 Pa and lower, the key comparison reference value p_R may be expressed in terms of the mean of measurement method means of <u>corrected</u> mean gauge readings ($p_j = f_C p_{jU}$) as follows:

$$p_R = \frac{1}{2}\left(\frac{1}{N_{MAN}} \sum_{j=1}^{N_{MAN}} p_j + \frac{1}{N_{DPB}} \sum_{j=1}^{N_{DPB}} p_j \right) = \frac{f_C}{2}\left(p_{MAN} + p_{DPB} \right) \qquad \text{(A1)}$$

where f_C is the required correction factor such that $p_R = p_t$, and the measurement method means of the <u>uncorrected</u> gauge readings for N_{MAN} liquid-column manometers, p_{MAN}, and for N_{DPB} double pressure balances[16], p_{DPB}, are calculated from

---

[15] The difference between method means could be due to either 1) unknown bias in one or both methods or 2) unknown procedural or instrumental bias not related to method at the laboratory using the method with a population of one. The mean of the method means is appropriate if 1) is true whereas the mean of the combined data may be more appropriate if 2) is true.

[16] The general case for $N_{DPB} > 1$ is developed here and then applied to the special case $N_{DPB} = 1$ for this comparison.

Table A1. Correction factors, f_C, reference values, p_R, and their estimated combined standard uncertainties, $u_c(p_R)$, calculated when excluding the result from MSL-NZ at 10 Pa (between dotted lines)†. The measurement method means of data obtained with liquid-column manometers, p_{MAN}, and with a double pressure balance, p_{DPB}, and estimates of their respective standard uncertainties, u_{MAN} and u_{DPB}, are also given. Not all digits are significant but are retained for calculation of final results.

Target Pressure Pa	N_{MAN}	p_{MAN} Pa	u_{MAN} Pa	N_{DPB}	p_{DPB} Pa	u_{DPB} Pa	f_C	p_R Pa	$u_c(p_R)$ Pa
1	2	1.0068	0.0046	1	0.9985	0.0065	0.99736	1	0.0040
3	3	3.0063	0.0032	1	3.0012	0.0086	0.99874	3	0.0046
10	3	10.0125	0.0036	0	10.0462†	0.0119	0.99876	10	0.0036
30	3	30.0256	0.0048	1	30.0191	0.0060	0.99926	30	0.0038
100	3	99.9735	0.0096	1	99.9664	0.0081	1.00030	100	0.0063
300	3	299.9720	0.0113	1	299.9937	0.0137	1.00006	300	0.0089
1000	3	999.9417	0.0208	1	999.9532	0.0191	1.000053	1000	0.014

† See footnote 18

$$p_{MAN} = \frac{1}{N_{MAN}} \sum_{j=1}^{N_{MAN}} p_{jU} \quad \text{and} \quad p_{DPB} = \frac{1}{N_{DPB}} \sum_{j=1}^{N_{DPB}} p_{jU} \tag{A2}$$

The uncertainties in the method means could be estimated from their sample variances

$$u_{MAN}^2 = \sum_{j=1}^{N_{MAN}} \frac{(p_{jU} - p_{MAN})^2}{N_{MAN}(N_{MAN}-1)} \quad \text{and} \quad u_{DPB}^2 = \sum_{j=1}^{N_{DPB}} \frac{(p_{jU} - p_{DPB})^2}{N_{DPB}(N_{DPB}-1)} \tag{A3}$$

if the "true" means of data taken with individual primary standards of a given method were equal without between-standard bias. This condition would be satisfied if primary standards of a given method were constructed to be exact replicates. However, primary standards at the participant laboratories were developed "in-house" and as such, each standard of a given method is unique even though they share a common basic operating principle. The Youden plots in Section 7.1 indicate that there is some relative bias between individual primary standards of a given method (only manometry is represented by more than one standard in this comparison). Equation (A3) would also be applicable if the number of standards of a given method were sufficiently large so that any between-standard biases could be regarded as being randomly distributed. But in this comparison, the sample size is too limited (1 to 3 depending on method and target pressure) and clearly Equation (A3) has no meaning for a sample size of one.

Alternatively, the uncertainties in p_{MAN} and p_{DPB} may be regarded as arising from the propagation of uncertainties $u_c(p_{jU})$ associated with independent values of p_{jU} in which case

$$u_{MAN}^2 = \sum_{j=1}^{N_{MAN}} \frac{u_c^2(p_{jU})}{N_{MAN}^2} \quad \text{and} \quad u_{DPB}^2 = \sum_{j=1}^{N_{DPB}} \frac{u_c^2(p_{jU})}{N_{DPB}^2} \tag{A4}$$

In the present comparison, $N_{DPB}=1$ and so $u_{DPB} = u_c(p_{jU})$ for the double pressure balance of MSL-NZ. Table A1 presents values for p_{MAN} and p_{DPB} and their uncertainties, which were calculated from Equations (A2), (16), and (A4) using data in Table 5.

Using the approximation $f_C \simeq 1$ (see Table A1), the combined uncertainty in p_R can then be estimated from:

$$u_c^2(p_R) = (f_c/2)^2 (u_{MAN}^2 + u_{DPB}^2) \simeq (1/2)^2 (u_{MAN}^2 + u_{DPB}^2) \tag{A5}$$

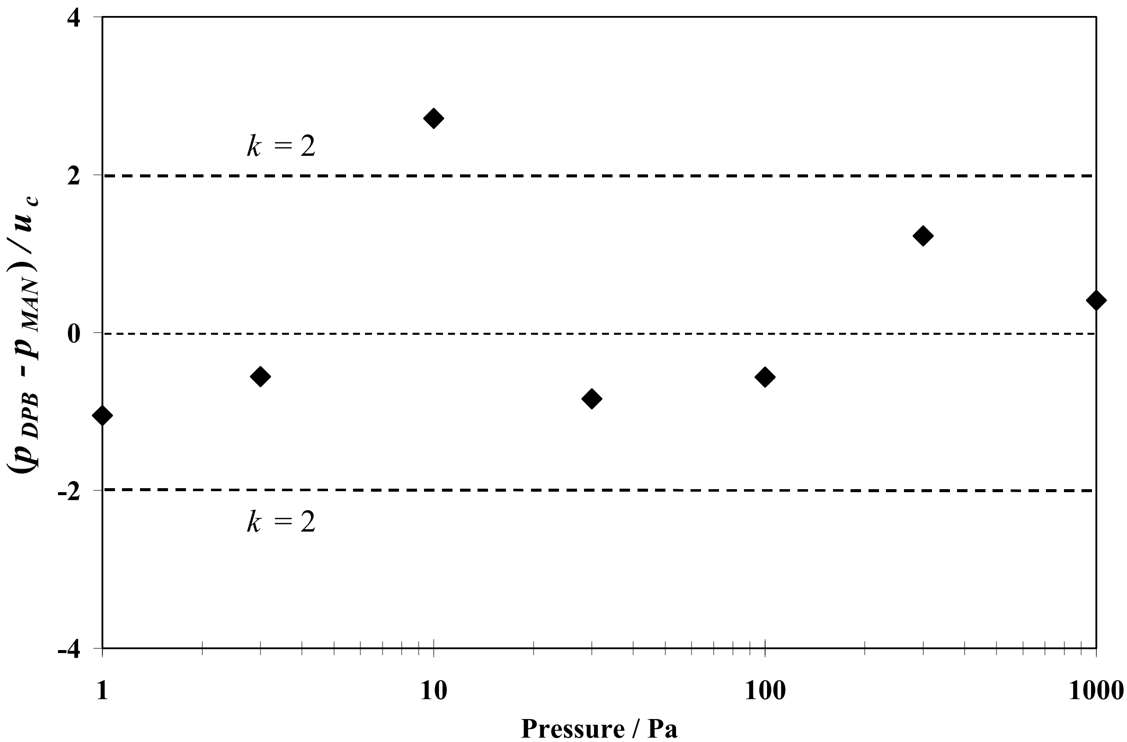


Figure A1. The ratio of the difference between measurement method means, $p_{DPB} - p_{MAN}$, and the standard uncertainty of this difference as a function of pressure.

where the uncertainties of the method means are given by Equation (A4) and the correction factor that sets the reference value equal to the target pressure is defined by Equation (A1) as

$$f_C \equiv 2p_t / \left(p_{MAN} + p_{DPB}\right) \qquad \text{(A6)}$$

Implicit in Equation (A5) is the assumption that any between-method bias is small and can be neglected. Between-method bias can arise when there are unknown systematic effects (biases) in one or both measurement methods that have not been taken into account in the stated uncertainties for primary standards of a given method. If the between-method bias is significant, its contribution to uncertainty in the reference value (i.e., the mean of the method means) can be estimated by a Type B evaluation (e.g., [14]) and added in quadrature to uncertainties of the method means in Equation (A5).

In order to check this assumption it is instructive to compare the difference between measurement method means, p_{DPB} - p_{MAN} , with its combined uncertainty, which can be estimated from

$$u_c^2 = u_{DPB}^2 + u_{MAN}^2 \qquad \text{(A7)}$$

Figure A1 presents the ratio, $(p_{DPB} - p_{MAN}) / u_c$, as a function of pressure. The plot shows that the difference between measurement method means[17] at 10 Pa lies outside two times the combined standard uncertainty of this difference. This is attributed to the result from MSL-NZ being moderately discrepant while being the only result for the double pressure balance method. To avoid undue influence, this result was not included in the calculation[18] of f_C, p_R and $u_c(p_R)$. At other pressures the difference between

[17] Actually the difference between the result from the double pressure balance of MSL-NZ and the manometer method mean.

[18] If included, the result from MSL-NZ at 10 Pa would continue to not satisfy the condition $|D_j| \leq U_j$ while changing f_c and $u_c(p_R)$ sufficiently so that results from two of the remaining three laboratories would no longer satisfy this condition (see Table 6).

methods is well within the $k = 2$ level indicating that the relative bias between the methods is rather small. For this reason its contribution to uncertainty in the reference value was neglected[19].

Table A1 presents the values for the correction factors, the reference values and their combined uncertainties for the case in which the result from MSL-NZ at 10 Pa has been excluded from the calculations.

A2. Calculation of Effective Degrees of Freedom and Associated Coverage Factors

In the present comparison, a conventional procedure described in references [10,11] was used to calculate coverage factors k_{95} that provide uncertainty intervals $U_j = k_{95}\, u_c(D_j)$ and $U_{jj'} = k_{95}\, u_c(D_{jj'})$, as defined in Equations (29) and (31), with a level of confidence approximating 95 %. The first step was to estimate the "effective degrees of freedom" ν_{eff} for the combined standard uncertainties $u_c(D_j)$ and $u_c(D_{jj'})$. Values of ν_{eff} were estimated by combining the degrees of freedom of individual component uncertainties using the Welch-Satterthwaite formula. The second step was to obtain the t-factor, $t_{95}(\nu_{eff})$, by interpolating a table of values given in the above cited references and then take $k_{95} = t_{95}(\nu_{eff})$.

The degrees of freedom of a component uncertainty obtained from a Type A evaluation can be readily determined by appropriate statistical methods. In the case of the uncertainties due to short-term random effects [see Equations (19) and (22)] the number of degrees of freedom is $\nu_{rdm} = 5 - 1 = 4$.

The degrees of freedom to associate with a component uncertainty obtained from a Type B evaluation is more problematic. However, if the component uncertainties are chosen in such a way that the probability of the measurand lying outside these limits is extremely small (e.g., when uncertainties are obtained from a rectangular probability distribution), then the degrees of freedom become infinitely large. This approximation, which cannot be fully justified for the estimates of $u_{std}(p_{ij})$ and $u_{lts}(p_{ij})$ in this comparison, is not necessarily unrealistic since the Type B evaluations were generally carried out in a manner that attempted to avoid an underestimation of the component uncertainties.

In the approximation that $\nu_{std} \to \infty$ and $\nu_{lts} \to \infty$, the effective degrees of freedom of the combined standard uncertainty $u_c(D_j)$ associated with the deviation of a primary standard from the KCRV was estimated from the Welch-Satterthwaite formula as:

$$\nu_{eff} = \frac{u_c^4(D_j)\nu_{rdm}}{\sum_{i=1}^{2}\left(1-\frac{1}{N}\right)^2 c^4 u_{rdm}^4(p_{ij}) + \sum_{j'=1}^{N_{MAN}}\sum_{i=1}^{2}\frac{c^4}{16N_{MAN}^4}u_{rdm}^4(p_{ij'}) + \sum_{j'=1}^{N_{DPB}}\sum_{i=1}^{2}\frac{c^4}{16N_{DPB}^4}u_{rdm}^4(p_{ij'})} \qquad \text{(A8)}$$

where N is either N_{MAN} or N_{DPB} depending on whether primary standard j is a manometer or a double pressure balance, and c is the (common) value for partial derivatives as defined in Equations (26) and (27). When the primary standard index j or j' refers to the pilot laboratory, the appropriate terms in the denominator must also include summations over the calibration number n and package label m as in Equation (27).

Similarly, the effective degrees of freedom of the combined standard uncertainty $u_c(D_{jj'})$ associated with the difference between primary standards was estimated from:

$$\nu_{eff} = \frac{u_c^4(D_{jj'})\nu_{rdm}}{\sum_{i=1}^{2} c^4 u_{rdm}^4(p_{ij}) + \sum_{i=1}^{2} c^4 u_{rdm}^4(p_{ij'})} \qquad \text{(A9)}$$

[19] Since there is no significant relative bias between methods, the reference values (actually the f_C needed to produce the nominal reference values) could also be calculated as the mean of the combined data. When the latter procedure is used there are only small changes in the values of D_j and U_j in Table 6 and the degrees of equivalence of the NMIs remain essentially unchanged.

where conditions affecting the summation of terms in the denominator of Equation (A8) apply equally well for Equation (A9).

Estimated values for the effective degrees of freedom and associated coverage factors k_{95}, which are needed to calculate expanded uncertainties $U_j = k_{95}\, u_c\,(D_j)$ and $U_{j\,j'} = k_{95}\, u_c\,(D_{j\,j'})$ in Table 6 of Section 7.2, are presented in Table A2.

Table A2. Estimates of the "effective degrees of freedom" ν_{eff} associated with standard uncertainties, $u_c\,(D_j)$ and $u_c\,(D_{j\,j'})$, and the coverage factors k_{95} that produce expanded uncertainties, $U_j = k_{95}\, u_c\,(D_j)$ and $U_{j\,j'} = k_{95}\, u_c\,(D_{j\,j'})$, with an approximate 95 % level of confidence.

				NMI$_{j'}$							
	Nominal			IMGC		MSL-NZ		NIST		NPL-UK	
NMI$_j$	Pressure	U_j		$U_{j\,j'}$		$U_{j\,j'}$		$U_{j\,j'}$		$U_{j\,j'}$	
	Pa	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}
IMGC	1	66	2.00			59	2.00	41	2.02		
	3	66	2.00			40	2.02	60	2.00	58	2.00
	10	140	2.00			28	2.04	118	1.98	115	1.98
	30	221	1.97			225	1.97	200	1.98	197	1.98
	100	556	1.97			554	1.97	502	1.97	221	1.97
	300	492	1.97			491	1.97	403	1.97	431	1.97
	1000	323	1.97			347	1.97	212	1.97	2037	1.96
MSL-NZ	1	40	2.02					79	1.99		
	3	15	2.13					16	2.12	13	2.16
	10	12	2.11					6	2.45	12	2.18
	30	125	2.04					1345	1.96	67	2.00
	100	393	1.97					13922	1.96	24	2.13
	300	294	1.98					1928	1.96	182	1.98
	1000	2296	1.96					6826	1.96	7002	1.96
NIST	1	44	2.01								
	3	21	2.09							610	1.97
	10	4779	2.11							369	1.97
	30	269	1.97							43	2.01
	100	450	1.97							16	2.12
	300	392	1.97							112	1.98
	1000	2581	1.96							6417	1.96
NPL-UK	3	19	2.09								
	10	738	2.12								
	30	68	2.00								
	100	23	2.08								
	300	164	1.98								
	1000	7507	1.96								

ACKNOWLEDGEMENTS

Contributions by several members of the Pressure and Vacuum Group at the NIST are gratefully acknowledged, in particular, S. A. Tison and C. R. Tilford, for their help with development of the transfer standard and with organization of the key comparison, F. G. Long and J. D. Kelley, for their construction of the two transfer standard packages, and S. Lu, for the development of data acquisition software and assistance with multiple calibrations of the transfer standards. The advice of C. R. Tilford throughout the course of this key comparison was also of great assistance. The many discussions with W. Guthrie of the Statistical Engineering Division at NIST concerning the statistical treatment of data were invaluable and are gratefully acknowledged as well.

REFERENCES

1. Comité International des Poids et Mesures (CIPM), *Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes*, 1999, Paris, pp.45.

2. Sutton C. M., "The accurate generation of small gauge pressures using twin pressure balances", *Metrologia*, 1986/87, **23**, 187-195.

3. Alasia F., Capelli A., Cignolo G., Goria R., Sardi M., "The MM1 laser interferometer low-range mercury manometer of the IMGC", *Metrologia*, 1999, **36** (6), 505-509.

4. Heydemann P. L. M., Tilford C. R., Hyland R. W., "Ultrasonic manometers for low and medium vacua under development at NBS", *J. Vac. Sci. Technol*., 1977, **14** (1), 597-605.

5. Tilford C. R., "Analytical procedure for determining lengths from fractional fringes", *Applied Optics*, 1977, **16** (7), 1857-1860.

6. Tilford C.R., Miiller A. P., "Development of a low differential-pressure standard", *Proc. 1997 NCSL Workshop and Symposium*, National Conference of Standards Laboratories, Boulder, CO, USA, 211-219.

7. Tilford C. R., "The speed of sound in a mercury ultrasonic interferometer manometer", *Metrologia*, 1987, **24**, 121-131.

8. Miiller A. P., "Measurement performance of capacitance diaphragm gauges and alternative low-pressure transducers", *Proc. 1997 NCSL Workshop and Symposium*, National Conference of Standards Laboratories, Boulder, CO, USA, 287-299.

9. Miiller A. P., "Measurement performance of high-accuracy low-pressure transducers", *Metrologia*, 1999, **36** (6), 617-621.

10. ISO, *Guide to the Expression of Uncertainty in Measurement* (International Organization for Standardization, Geneva, Switzerland, 1993).

11. Taylor B. N., Kuyatt C. E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, 1994 Edition (U.S. Government Printing Office, Washington, 1994).

12. Youden W. J., "Graphical diagnosis of interlaboratory test results", *Industrial Quality Control*, 1959, **15** (11), 24-28.

13. Guthrie W., Statistical Engineering Division, National Institute of Standards and Technology, *Private Communication.*

14. Levenson M. S., Banks D. L., Eberhardt K. R., Gill L. M., Guthrie W. F., Liu H. K., Vangel M. G., Yen J. H., Zhang N. F., “An approach to combining results from multiple methods motivated by the ISO GUM”, *J. Res. Natl. Inst. Stand. Technol*., 2000, **105** (4), 571-579.